THE BELL SYSTEM
TECHNICAL JOURNAL

VOLUME XXXV MAY 1956 NUMBER 3

Chemical Interactions Among Defects in Germanium and Silicon

By HOWARD REISS, C. S. FULLER, and F. J. MORIN

Interactions among defects in germanium and silicon have been investigated. The solid solutions involved bear a strong resemblance to aqueous solutions insofar as they represent media for chemical reactions. Such phenomena as acid-base neutralization, complex ion formation, and ion pairing, all take place. These phenomena, besides being of interest in themselves, are useful in studying the properties of the semiconductors in which they occur. The following article is a blend of theory and experiment, and describes developments in this field during the past few years.

CONTENTS

I. INTRODUCTION

The effort of Wagner[1] and his school to bring defects in solids into the domain of chemical reactants has provided a framework within which various abstruse statistical phenomena can be viewed in terms of the intuitive principle of mass action.[2] Most of the work to date in this field has been performed on oxide and sulfide semiconductors or on ionic compounds such as silver chloride. In these materials the control of defects (impurities are to be regarded as defects) is not all that might be desired, and so with a few exceptions, experiments have been either semiquantitative or even qualitative.

With the emergence of widespread interest in semi-conductors, culminating in the perfection of the transistor, quantities of extremely pure single crystal germanium and silicon have become available. In addition the physical properties, and even the quantum mechanical theory of the behavior of these substances have been widely investigated, so that a great deal of information concerning them exists. Coupled with the fact that defects in them, especially impurities, are particularly susceptible to control, these circumstances render germanium and silicon ideal substances in which to test many of the concepts associated with defect interactions.

This view was adopted at Bell Telephone Laboratories a few years ago when experimental work was first undertaken. Not only has it been possible to demonstrate quantitatively the validity of the mass action principle applied to defects, but new kinds of interactions have been discovered and studied. Furthermore new techniques of measurement have been developed which we feel open the way for broader investigation of a still largely unexplored field.

In fact solids (particularly semiconductors like germanium and silicon)

appear in every respect to provide a medium for chemical reactivity similar to liquids, particularly water. Such pehnomena as acid-base reactions, complex ion formation, and electrolyte phenomena such as Debye Hückel effects, ion pairing, etc., all seem to take place.

Besides the experiments theoretical work has been done in an attempt to define the limits of validity of the mass action principle, to furnish more refined electrolyte theories, and most importantly, to provide firm theoretical bases for entirely new phenomena such as ion pair relaxation processes.

The consequence is that the field of diamond lattice[3] semiconductors which has previously engaged the special interests of physicists threatens to become important to chemists. Semiconductor crystals are of interest, not only beeause of the specific chemical processes occurring in these substances, but also because they serve as proving grounds for eertain ideas current among chemists, such as electrolyte theory. On the other hand renewed interest is induced on the part of physicists because chemical effects like ion pairing engender new physical effects.

The purpose of this paper is to present the field of defect interaction as it now stands, in a manner intelligible to both physicists and chemists. However, this is not a review paper. Most of the experimental results, and particularly the theories which are fully derived in the text or the appendices are entirely new. Some allusion will be made to published work, particularly to descriptions of the results of some previous theories, in order to round out the development.

The governing theme of tbe article lies in the analogy between semiconductors and aqueous solutions. This analogy is useful not so much for what it explains, but for the experiments which it suggests. More than once it has stimulated us to new investigations.

In our work we have made extensive use of lithium as an impurity. This is so because lithium can be employed with special ease to demonstrate most of the eoncepts we have in mind. This specialization should not obscure the fact that otber impurities although not well suited to the performance of accurate measurements, will exhibit much of the same behavior.

II. ELECTRONS AND HOLES AS CHEMICAL ENTITIES

Since electrons and holes[4] are obvious occupants of semiconductors like germanium and silicon, and are intimately associated with the presence of donor and acceptor impurities,[4] it is fitting to inquire into the roles they may play in chemical interactions between donors and ac-

ceptors. This question bas been discussed in two papers,[5, 6] and only its principle aspects will be considered.

To gain perspective it is convenient to consider a system representing the prototype of most systems to be discussed here. Consider a single crystal of silicon containing substitutional boron atoms. Boron, a group III element, is an acceptor, and being substitutional cannot readily diffuse[7] at temperatures mucb below tbe melting point of silicon. If this crystal is immersed in a solution containing lithium, e.g., a solution of lithium in molten tin, lithium will diffuse into it and behave as a donor. Evidence suggests that lithium dissolves interstitially in silicon, thereby accounting for the fact that it possesses a high diffusivity[8] at a temperature where boron is immobile, for example, below 300°C. When the lithium is uniformly distributed throughout the silicon its solubility in relation to the external pbase can be determined. Throughout this process boron remains fixed in the lattice.

If both lithium and boron were inert impurities the solubility of the former would not be expected to depend on the presence or absence of the latter, for the level of solubility is low enough to render (under ordinary circumstances) the solid solution ideal.[9] On the other hand the impurities exhibit donor and acceptor behaviors respectively, and some unusual effects might exist. We shall first speculate on the simplest possibility in this direction, with tbe assistance of the set of equilibrium reactions diagrammed below.*

$$\begin{array}{ccccc} Li(Sn) & \rightleftarrows & Li(Si) & \rightleftarrows & Li^+ + e^- \\ & & & & + \\ B(Si) & \rightleftarrows & B^- & + & e^+ \\ & & & & \Updownarrow \\ & & & & e^+e^- \end{array} \qquad (2.1)$$

At the left litbium in tin is sbown as $Li(Sn)$. It is in reversible equilibrium with $Li(Si)$, un-ionized litbium dissolved in silicon. The latter, in turn, ionizes to yield a positive Li^+ ion and a conduction electron, e^-. Boron, confined to the silicon lattice as $B(Si)$ ionizes as an acceptor to give B^- and a positive hole, e^+. The conduction electron, e^-, may fall into a valence band bole, e^+, to form a recombined hole-electron pair, e^+e^-. This process and its reverse arc indicated by the vertical equilibrium at the rigbt.

All of the reactions in (2.1), occuring within the silicon crystal are describable in terms of tansitions between states in the energy band dia-

* A glossary of symbols is given at the end of this article.

gram of silicon, exhibited in Fig. 1. The conduction band, tbe valence band, and the forbidden gap are shown. Lithium and boron both introduce localized energy states in the range of forbidden energies. The state for lithium lies just below tbe bottom of the conduction band while that for boron lies just over the top of tbe valence band. The separations in energy between most donors or acceptors and their nearest bands are of the order of hundredths of an electron volt while the breadths of the forbidden gaps in germanium or silicon are of the order of one electron volt.

Process 1 in Fig. 1 involving a transition between the donor level and conduction band corresponds to the ionization of lithium in (2.1). Process 2 is the ionization of boron while process 3 represents hole-electron recombination and generation. The various energies of transition are the heats of reaction of the chemical-like changes in (2.1).

Proceeding in the chemists fashion one might argue as follows concerning (2.1). If e^+e^- is a stable compound, as it is at fairly low temperatures, then its formation should exhaust the solution of electrons, forcing tbe set of lithium equilibria to the right. In this way the presence of boron, supplying holes toward the formation of e^+e^-, increases the solubility of lithium. In fact if e^+ is regarded as the solid state analogue of the hydrogen ion in aqueous solution, and e^- as the counterpart of the hydroxyl ion, then the donor, lithium, may be considered a base while boron, may be considered an acid. Furthermore e^+e^- must correspond to water. Thus the scheme in (2.1) is analogous to a neutralization reaction in which the weakly ionized substance is e^+e^-.

If the immobile boron atoms were replaced by immobile donors, e.g., phosphorus atoms, a reduction, rather than an increase, in the solubility

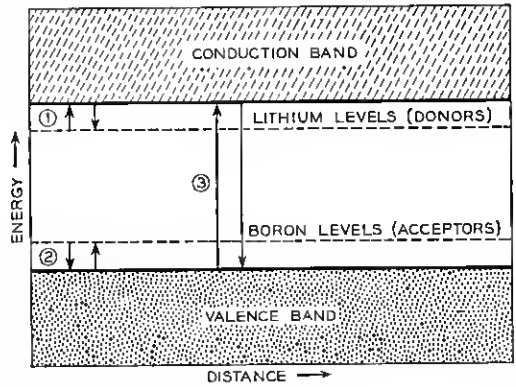

Fig. 1 — Energy band diagram showing the chemical equilibria of (2.1).

of lithium might be expected on the basis of an oversupply of electrons (i.e., by the common ion effect[10]). In that case we would have a base displacing another base from solution.

The intimate comparison between this kind of solution and an aqueous solution is worth emphasizing not so much for what it adds to one's understanding of the situation but rather for the further effects it suggests along the lines of analogy. These additional phenomena have been looked for and found, and will be discussed later in this article.

The scheme shown in (2.1) should be applicable, in principle, to other donors and acceptors and to germanium and other semiconductors as well as silicon. Furthermore the external phase may be any one of a suitable variety, and need not even be liquid. Other systems, however, are not as convenient, especially in regard to the ease of equilibration of an impurity over the parts of an heterogeneous system. The lengths to which one can go in comparing electrolytes and semiconductors are discussed in a recent paper.[11]

In order to quantify the scheme of (2.1) it seems natural to invoke the law of mass action.[2] Treatments in which holes and electrons are involved in mass action expressions are not new, although systems forming such perfect analogies to aqueous solutions do not seem to have been discussed in the past. For example, in connection with the oxidation of copper Wagner[12] writes

$$4Cu + O_2 \rightleftarrows 2Cu_2O + 4\square^{-} + 4e^{+} \tag{2.2}$$

in which $\square^{-}$ is a negatively charged cation vacancy in the Cu_2O lattice, and e^{+} is a hole. Wagner proceeds to invoke the law of mass action in order to compute the oxygen pressure dependence in this system.

In another example Baumbach and Wagner[13] and others have investigated oxygen pressure over non-stoichiometric zinc oxide. They consider the possible reactions

$$\begin{array}{l} 2ZnO \rightleftarrows 2Zn + O_2 \\ \qquad\quad \uparrow\downarrow \\ \qquad\quad 2Zn^{+} \rightleftarrows 2Zn^{++} + 2e^{-} \\ \qquad\quad + \\ \qquad\quad 2e^{-} \end{array} \tag{2.3}$$

and apply the law of mass action. In (2.3) the various states of Zn are presumably interstitial.

Kroger and Vink[14] have recently considered the problem in oxides and sulfides in a rather general way. However in none of the oxide-sulfide systems has it been possible to achieve really quantitative results. In

contrast silicon and germanium offer possihilities of an entirely new order. The advent of the transistor has not only provided large supplies of pure single crystal material, hut it has also made availahle a store of fundamental information concerning the physical properties of these suhstances. For example, data exists on their energy band diagrams including impurity states — also on resistivity — impurity density curves, diffusivities of impurities, etc. Furthermore, the amount of ionizable impurities can he controlled within narrow limits, and can he changed at will and measured accurately. Consequently it is reasonahle to assume that experiments on germanium and silicon will be more successful than similar investigations using other materials.

At this point it is in order to examine whetber or not the treatment of electrons and holes as normal chemical entities satisfying the law of mass action is altogetber simple and straightforward. This problem has been investigated by Reiss[5] who found the treatment permissible only as long as the statistics satisfied hy holes and electrons remain classical. The validity of this contention can he seen in a very simple manner. Consider a system like that in (2.1). Let the total concentration of donor (ionized and un-ionized) he N_D, the concentration of ionized donor he D^+, the concentration of conduction electrons he n, and that of valence band holes he p. Let N_A and A^- denote the concentrations of total acceptor and acceptor ions respectively. Finally, let α he the thermodynamic activity[15] of the donor (lithium in (2.1)) in the external phase.

Theu, corresponding to the heterogeneous equilihrium in which lithium distributes itself hetween the two phases we can write

$$\frac{N_D - D^+}{\alpha} = K_0 \tag{2.4}$$

in which K_0 depends on temperature, but not on composition. This assumes the semiconductor to he dilute enough in donor so that the activity of un-ionized donor can be replaced hy its concentration, $N_D - D^+$. For the ionization of the donor we can write the mass action relation,

$$\frac{D^+ n}{N_D - D^+} = K_D \tag{2.5}$$

and for tbe acceptor,

$$\frac{A^- p}{N_A - A^-} = K_A \tag{2.6}$$

while for the electron-hole recomhination equilibrium

$$np = K_1 \tag{2.7}$$

In (2.5), (2.6), and (2.7) all the K's are independent of composition. To these equations is added the charge neutrality condition,

$$D^+ + p = A^- + n \tag{2.8}$$

Equations (2.4) through (2.8) are enough to determine N_D in its dependence on N_A, α, and the various K's. Together they represent the mass action approach. To demonstrate their validity it is necessary to appeal to statistical considerations.

Thus $N_D - D^+$, the concentration of un-ionized donor is really the density of electrons in the donor level of the energy diagram for the semiconductor. According to Fermi statistics this density is given by[5]

$$N_D - D^+ = N_D/\{1 + \tfrac{1}{2} \exp [(E_D - F)/kT]\} \tag{2.9}$$

in wbich E_D is the energy of the donor level, F is the Fermi level,[16] k, the Boltzmann constant, and T, the temperature. Furthermore, according to Fermi statistics, n, the total density of electrons in the conduction band is

$$n = \sum_i g_i/\{1 + \exp [(E_i - F)/kT]\} \tag{2.10}$$

where g_i is the density of levels of energy, E_i, in the conduction band, and the sum extends over all states in that band. Similar expressions are available for the occupation of the acceptor level and the valence band. F is usually determined by summing over all expressions like (2.9) and (2.10) and equating the result to the total number of electrons in the system. This operation corresponds exactly to applying the conservation condition, (2.8). It is obvious from the manner of its determination tbat F depends upon $N_D - D^+$, n, etc.

If we now form the expression on the left of (2.5) by substituting for each factor in it from (2.9) and (2.10), it is obvious that the result depends in a very complicated fashion upon F, and so cannot be the constant, K_D, independent of composition, since in the last paragraph F was shown to depend on composition. On the other hand if attention is confined to the limit in which classical statistics apply[17] the unities in the denominators of (2.9) and (2.10) can be disregarded in comparison to the exponentials, and those equations become

$$N_D - D^+ = 2N_D e^{F/kT} e^{-E_D/kT} \tag{2.11}$$

and

$$n = e^{F/kT} \sum g_i e^{-E_i/kT} \tag{2.12}$$

respectively. Moreover, from (2.11)

$$D^+ = N_D[1 - 2e^{F/kT}e^{-E_D/kT}] = N_D \tag{2.13}$$

where the second term in brackets is ignored for the same reason as unity in the denominators of (2.9) and (2.10). Substituting (2.11) through (2.13) into (2.5) yields

$$\frac{D^+n}{N_D - D^+} = \frac{\sum_i g_i e^{-E_i/kT}}{2e^{-E_D/kT}} \tag{2.14}$$

in which the right side is truly independent of composition, since F has cancelled out of the expression. Similar arguments hold for (2.6) and (2.7). Therefore in the classical limit the law of mass action is valid, at least insofar as internal equilibria are concerned.

We have next to examine the validity of (2.4) which is really the law of mass action applied to the heterogeneous equilibrium between phases. Substitution of (2.11) into (2.4) leads to the prediction

$$\alpha = \frac{2e^{-E_D/kT}}{K_0}\{e^{F/kT}\}N_D = K\{e^{F/kT}\}N_D \tag{2.15}$$

in the classical case, if (2.4) is valid. In order to confirm (2.15) it is necessary to evaluate the chemical potentials[18] of the donor in the external phase and in the semiconductor, and equate the two. The resulting expression should be equivalent to (2.15).

Since α is the activity of the donor in the external phase its chemical potential in that phase is, by definition,

$$\mu = \mu^0(T, p) + kT\,\ell n\,\alpha \tag{2.16}$$

where μ^0, the chemical potential in the standard state, may depend on temperature and pressure, but not on composition. To compute the chemical potential in the semiconductor statistical methods must once more be invoked. Thus, according to (2.13), donor atoms are nearly totally ionized in the classical case, so that the addition of a donor atom to the semiconductor amounts to addition of two separate particles, the donor ion and the electron. The chemical potential of the added atom is therefore the sum of the potentials of the ion and the electron separately. Since the ions are supposedly present in low concentration the latter can serve as an activity,[19] and in analogy to (2.16) we obtain for the ionic chemical potential

$$\mu_{D^+} = \mu_{D^+}{}^0(T, p) + kT\,\ell n\,D^+ \tag{2.17}$$

Furthermore, it is well established[20] that the Fermi level plays the role of chemical potential, μ_e, for the electron

$$\mu_e = F \tag{2.18}$$

Thus the chemical potential for the donor atom is

$$\begin{aligned} \mu_D &= \mu_{D^+} + \mu_e = \mu_{D^+}{}^0 + kT\,\ell n D^+ + F \\ &= \mu_{D^+}{}^0 + kT\,\ell n\, N_D + F = \mu_{D^+}{}^0 + kT\,\ell n\,\{e^{F/kT}\}N_D \end{aligned} \tag{2.19}$$

where (2.13) has been used to replace D^+ by N_D. We note that the activity of the donor atom must be

$$\{e^{F/kT}\}N_D \tag{2.20}$$

with $e^{F/kT}$ playing the role of an activity coefficient.[21]

Equating μ_D given by (2.19) to μ in (2.16) results in the equation

$$\alpha = \exp[(\mu_{D^+}{}^0 - \mu^0)/kT]\{e^{F/kT}\}N_D \tag{2.21}$$

which can be made identical to (2.15) by identifying

$$\exp[(\mu_{D^+}{}^0 - \mu^0)/kT]$$

with K of that expression. Thus in the classical case the law of mass action is applicable to the heterogeneous equilibrium.

When classical statistics no longer apply it is still possible to evaluate $N_D - D^+$, using the full expression (2.9). Therefore the solubility N_D, of the donor can still be determined if (2.4) remains valid. To decide this question it is necessary to evaluate μ_D, the chemical potential of the donor in the semiconductor under non-classical conditions. This problem is not as simple as those treated above, but it can be solved, and the detailed arguments can be found in Reference 5. Here we shall be content with quoting the results. However, before doing this the non-classical counterpart of (2.15) will be written by combining (2.9) with (2.4). The result is

$$\alpha = [K_0/\{1 + \tfrac{1}{2}\exp[(E_D - F)/kT]\}]N_D \tag{2.22}$$

and if (2.4) is valid (2.22) should be derivable by equating μ to the proper value of μ_D.

Since in the non-classical case a finite portion of the donor states are occupied by electrons, the introduction of an additional *average* donor atom is no longer equivalent to adding two independent particles whose chemical potentials can be summed. In the statistical derivation of μ_D it is therefore necessary to evaluate the total free energy of the semi-

conductor phase, and to differentiate this with respect to N_D, keeping temperature and pressure fixed.* The result is

$$\mu_D = \mu_{D^+}{}^0 + kT \ln N_D + F - kT \ln \{1 + 2 \exp[-(E_D - F)/kT]\} \tag{2.23}$$

in which it has been assumed that the concentration of impurity is sufficiently low so that the solution would be ideal if the impurity could not ionize. In the classical case the exponential in the logarithm is small compared to unity and (2.23) becomes identical with (2.19), as it should. In the totally degenerate case the exponential dominates tbe unity and we have

$$\begin{aligned} \mu_D &= \{\mu_{D^+}{}^0 + E_D - kT \ln 2\} + kT \ln N_D \\ &= \mu_D^0 + kT \ln N_D \end{aligned} \tag{2.24}$$

which is the chemical potential of an un-ionized component of a dilute

* An interesting by-product of this derivation (discussed in Reference 5) is the fact that the Fermi level, F, is hardly ever the Gibbs free energy per electron for the electron assembly, although it is always the electronic chemical potential, in the sense that it measures the direction of flow of electrons. This arises because the Gibbs free energy is not always a homogeneous function[22] of the first degree in the mole numbers (electron numbers). Thus if the number of electrons in the assembly is N, the Gibbs free energy, G, is given by

$$G = NF + kT \sum_j \left[V \left\{ \frac{\partial \omega_j}{\partial V} \right\}_{T,N} - \omega_j \right] \ln \frac{\omega_j}{h_j}$$

where the sum is over all energy levels, j, referred to an invariant standard level. V is the volume of the system, ω_j is the total number of states at the jth level, and h_j is the number of unoccupied states (holes) at the jth level. For F to be the free energy per electron the term involving the sum must vanish so that

$$F = \frac{G}{N}$$

But this can only happen when

$$\omega_j = K_j V$$

where K_j is independent of V. This requirement is formally met in the case of the free electron gas where the electrons have been treated as independent particles in a box so that

$$\omega_j = [8m_0{}^{3/2} \pi E \, dE/2h^3]V$$

where m_0 is the electron mass, and h, Plank's constant. Since this is the case most frequently dealt with in thermodynamic problems it has been customary to think of F as the free energy per electron, although even here the truth of the contention depends on the assumption of particle in the box behavior.

At the other extreme, it is obvious that ω_j for a level corresponding to the deep closed shell states of the atoms forming a solid cannot depend at all on the external volume since they are essentially localized. In computing the free energy of the semiconductor phase it is necessary to understand carefully subtleties of this nature.

solution, as it should be for the degenerate case in which ionization is suppressed. Equating μ_D in (2.23) to μ in (2.16) yields

$$\alpha = \left\{ \frac{\frac{1}{2} \exp\left[(\mu_{D^+}{}^0 - \mu^0 + E_D)/kT\right]}{1 + \frac{1}{2} \exp\left[(E - F)/kT\right]} \right\} N_D \tag{2.25}$$

which is identical with (2.22) if K_0 is taken to be

$$\frac{1}{2} \exp[(\mu_{D^+}{}^0 - \mu^0 + E_D)/kT] \tag{2.26}$$

Thus one arrives at the conclusion that the law of mass action remains valid for the heterogeneous equilibrium even when it fails for the homogeneous internal equilibria.

This is a fairly important result since it implies that solubilities can give information on the behavior of the Fermi level and hence on the distribution of electronic energy levels, even under conditions of degeneracy.

The chemical potential specified by (2.23) is of course important in itself, for treating any equilibrium (external or internal) in which the donor may participate.

One last remark is in order. This concerns the treatment of heterogeneous equilibria involving some external phase, and the surface[23] rather than the body of a semiconductor. In such treatments it has been customary to compute the chemical potential of an ionizable adsorbed atom by summing the ion chemical potential and the Fermi level, as in (2.19). This is no more possible if the statistics of the surface states are non-classical, then it is possible when considering non-classical situations involving the body of the crystal. Care must therefore be exercised also in the treatment of surface equilibria.

The above discussion has shown that there are extensive ranges of conditions under which holes and electrons obey the law of mass action, and behave like chemical entities. In the next section some of the consequences of this fact will be developed.

III. APPLICATION OF THE MASS ACTION PRINCIPLE

Equations (2.4) through (2.8) will now be used to determine how, in the classical case, the solubility, N_D, of lithium in (2.1) depends upon N_A the concentration of boron in silicon. In the experiments to be described, the systems are classical, and the donors and acceptors therefore so thoroughly ionized that N_D can be replaced by D^+ and N_A by A^-. Insertion of (2.4) into (2.5) yields

$$D^+ n = \alpha K_D K_0 = K^* \tag{3.1}$$

since α is maintained constant. Furthermore (2.7) can be written as

$$np = K_1 = n_i^2 \tag{3.2}$$

where n_i is obviously the concentration of holes or electrons under the condition that the two are equal. It is called the intrinsic concentration[24] of holes or electrons. The values of n_i in germanium and silicon have been determined by Morin.[25, 26] Fig. 2 gives plots of the logarithms of n_i in germanium and silicon versus the reciprocals of temperature. These results are necessary for subsequent calculations.

Since N_A and A^- are assumed equal, we may dispense with (2.6). The one remaining equation is then (2.8) which we adopt unchanged. These three relations, (3.1), (3.2), and (2.8) are sufficient to determine D^+ or N_D as a function of A^- or N_A. The only undetermined parameter in the set is K^* and this can be evaluated by measuring the solubility, D^+, in the absence of acceptor, i.e., under the condition that A^- is zero. The symbol $D_0{}^+$ is used to designate this value of D^+. In Reference 6 it is shown that

$$D_0{}^+ = K^*/(K^* + n_i^2)^{1/2}$$

or

$$K^* = (D_0{}^+)^2/2 + \{(D_0{}^+)^4/4 + n_i^2(D_0{}^+)^2\}^{1/2} \tag{3.3}$$

Eliminating K^* by the use of this relation it is further shown in Reference 6 that

$$D^+ = \frac{A^-}{1 + \sqrt{1 + (2n_i/D_0{}^+)^2}} + \left\{\left[\frac{A^-}{1 + \sqrt{1 + (2n_i/D_0{}^+)^2}}\right]^2 + (D_0{}^+)^2\right\}^{1/2} \tag{3.4}$$

which is the required relation between donor solubility and acceptor concentration.

Examination of (3.4) reveals several simple features, the more important of which we list below:

(1) When A^- (the acceptor doping) is sufficiently large so that $(D_0{}^+)^2$ in the second term can be ignored relative to the term in A^-, (3.4) reduces to that of a straight line with slope

$$D^+/A^- = \frac{2}{1 + \sqrt{1 + (2n_i/D_0{}^+)^2}} \tag{3.5}$$

Knowledge of this slope is equivalent to knowledge of $D_0{}^+$.

(2) Where the straight line portion of the D^+ versus A^- curve is in-

volved, the temperature dependence of the solubility, D^+, enters only through the ratio, $n_i/{D_0}^+$. If this ratio is very small, then

$$D^+ \approx A^- \tag{3.6}$$

and the solubility is independent of temperature. In this condition D^+ may approximate A^- hy being either slightly less or slightly greater than the latter. Details are given in Reference 6.

(3) Whereas D^+ at small values of doping may he an increasing function of temperature, it may, depending on the system, he a decreasing function of temperature at high dopings. Thus doping may change the sign of the temperature coefficient of soluhility. Because of this, doping sometimes may prevent precipitation of a donor when a semiconductor is cooled, since the latter hecomes an undersaturated rather than a supersaturated solution of impurity. Details are given in Reference 6.

(4) It is also shown in Reference 6 that for the acceptor to have any effect on the soluhility of the donor the concentration of A^- should satisfy the following criterion

$$A^- > ({D_0}^+ \quad \text{or} \quad n_i) \tag{3.7}$$

${D_0}^+$ or n_i heing used depending on which is greater. Ohviously at high

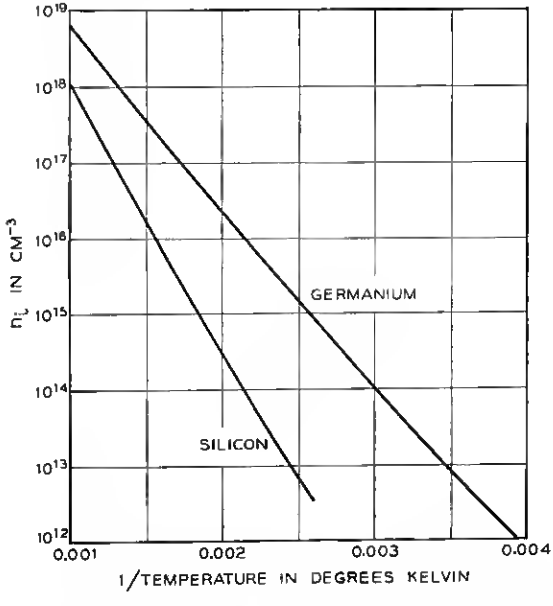

Fig. 2 — Temperature dependences of intrinsic carrier concentrations in germanium and silicon.

temperatures when n_i achieves a very large value it may not be possible to have A^- exceed n_i, and no effect due to the acceptor will be observable. This is simply a mathematical reflection of the fact that tbe hypothetical compound e^+e^- in (2.1) is highly dissociated at high temperatures so that the holes contributed by the acceptor cannot cause the exhaustion of electrons in the solution.

In Reference 6 the system described in (2.1) was investigated for the purpose of testing (3.4). The concentrations, D^+ and A^-, of lithium and boron respectively were determined by measuring the electrical resistivities of the crystal specimens before and after immersion in molten tin contaning lithium. Some typical results of these experiments are shown iu Fig. 3 which contains three D^+ versus A^- isotherms for the temperatures 249°, 310°, and 404°C. For the case shown the tin phase contained 0.18 per cent lithium by weight.

The points in the figure represent experimeutal findings, while the drawn curves are based on theory. The agreement between theory and experiment is very good, in fact the overall accuracy appears to be better than 1 per cent. These isotherms are only a few of a large group obtained at different temperatures and with differently proportioned external phases. The accuracy in all of these is of the same order.

Various of the features of (3.4) listed above are apparent in the curves of Fig. 3. For example at large values of A^- the curves are straight lines, thus validating (3.5). Also, the inversion of the temperature coefficieut of solubility with doping is apparent for the curves cross one another, and whereas, at low dopings (low A^-) the solubility is an increasing function of temperature, at high dopings it decreases with increasing temperature. Finally we note that D^+ remains more or less independent of A^- until A^- exceeds n_i, confirming (3.7). Values of n_i appear in the Figure.

The possible increases in solubility above ${D_0}^+$ are really quite large. For example in Fig. 3 the largest increase is of the order of a factor of 10^3. However in some experiments increases of 10^6 have been observed. These effects truly represent profound interactions between impurities which are present in highly attenuated form. Thus the number of atoms per cubic centimeter in crystal silicon is of the order of 5×10^{22} cm^{-3}. Interactions at doping levels as low as 10^{15} cm^{-3}, as appear in Fig. 3, therefore take place at atom fraction levels of about 2×10^{-8}.

In Fig. 4 we show a curve of lithium solubility at room temperature in gallium-doped germanium. The curve is wholly experimental; no attempt has been made to apply theory. The symbols D^+ and A^- are once more used for the donor and acceptor. In this case the curve again exhibits some of the general features required by (3.4). The measure-

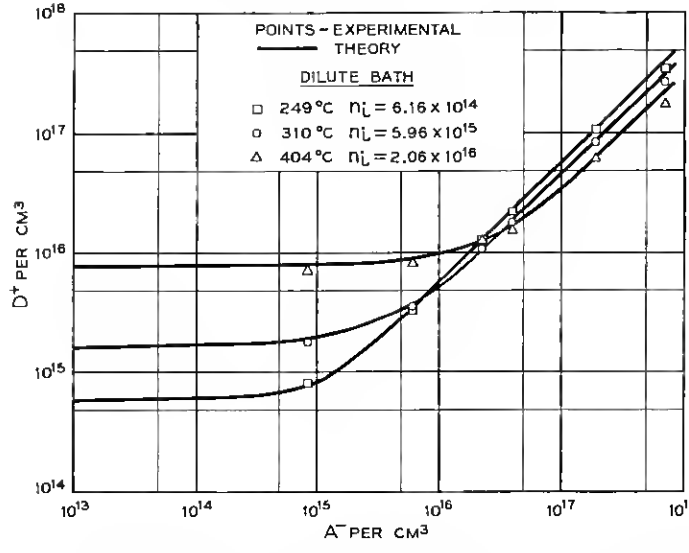

Fig. 3 — Isotherms showing the solubility of lithium D^+, in silicon as a function of boron doping A^-, for an external phase of tin containing 0.18 per cent lithium.

ments were made by saturating gallium-doped germanium crystals with lithium by alloying lithium to the germanium surface at a high temperature, and letting it diffuse in. Following this the crystals were cooled and lithium was allowed to precipitate to equilibrium. In this case the external solution is the precipitate and is of unknown composition.

If the straight line portion of the curve is used to determine D^+/A^- appearing in (3.5), the value of $D_0{}^+$ associated with the precipitate as an external phase can be computed by using the value of n_i obtained from Fig. 2 for 25°C. The latter is 3×10^{13} cm^{-3}, and the measured D^+/A^- is 0.85. Application of (3.5) then leads to a value of $D_0{}^+$ of 6.6×10^{13} cm^{-3} at 25°C. Since the highest value of D^+ measured in Fig. 4 is 5.5×10^{18} cm^{-3}, tbe solubility increase here shows a factor of 10^5. Interaction is already apparent at values of A^- as low as 10^{14} cm^{-3}, and since there are 4.4×10^{22} cm^{-3} atoms per cubic centimeter in pure germanium this represents interaction at levels of atom fraction as low as 2×10^{-9}.

IV. FURTHER APPLICATIONS OF THE MASS ACTION PRINCIPLE

In the last section the possibility was mentioned of inverting the sign of the temperature coefficient of solubility, and so preventing impurity

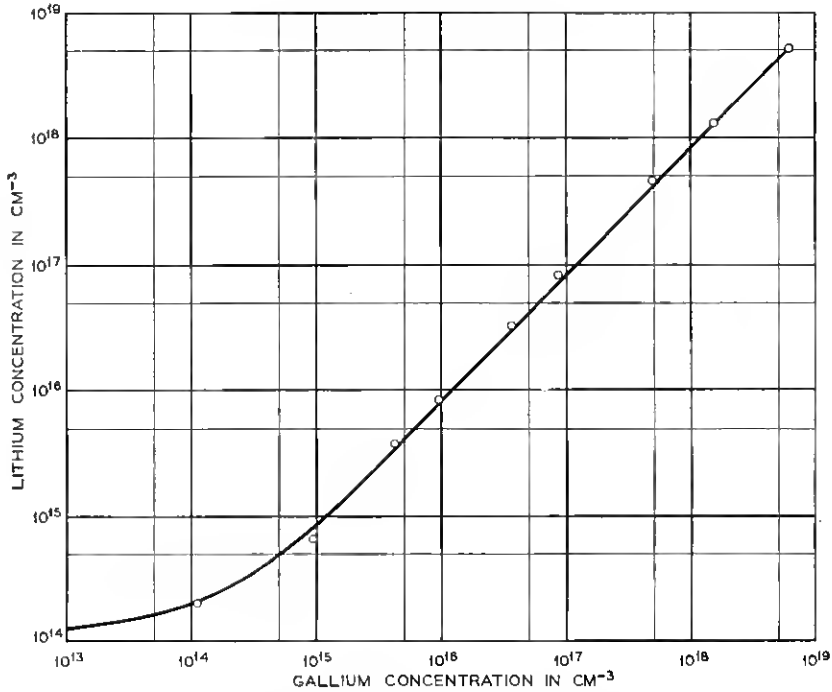

Fig. 4 — Room temperature isotherm showing the solubility of lithium in germanium as a function of gallium doping, the external phase being an alloy of lithium and germanium. The curve merely shows locus of experimental points.

precipitation which might normally occur upon cooling a crystal specimen. An experiment demonstrating this effect is described in Reference 6. Two specimens of germanium, one without added acceptor, and the other containing gallium at an estimated concentration of 1.3×10^{19} cm^{-3}, were saturated with lithium. Table I compares the changes in lithium content observed in these samples with the passage of time. After 25 days no apparent precipitation had occurred in the gallium doped specimen, while precipitation was almost complete in the other.

This result suggests a practical scheme for measuring the concentration of lithium along the solidus curve of the lithium-germanium phase diagram, i.e., the solubility of lithium in solid germanium when the external phase is also composed of germanium and lithium, and probably represents the liquidus phase. This measurement, though desirable, has not been performed before because lithium, diffused into germanium at an elevated temperature, precipitates when the specimen is cooled.

TABLE I

Ga Conc. (cm^{-3})	Li Conc. after saturation (cm^{-3})	Li Conc. after 4 days at room Temp. (cm^{-3})	Li Conc. after 25 days at room Temp. (cm^{-3})
0	1.4×10^{15}	9.0×10^{15}	1.1×10^{15}
1.3×10^{19}	8.0×10^{18}	8.0×10^{18}	8.0×10^{18}

Resistivities then measure only the dissolved lithium although the true solubility at the temperature of saturation includes the precipitated material.

However, we have seen that germanium suitably doped with gallium will not lose lithium by precipitation. Therefore the experiment might be performed in doped germanium. The only difficulty with this suggestion lies in the fact that *doping changes the solubility*. This objection can be overcome through use of (3.4). In terms of that equation D^+ would be measured in the presence of gallium whereas $D_0{}^+$, the solubility in undoped germanium, is required. But according to (3.4) if D^+, n_i, and A^- (gallium concentration) are known $D_0{}^+$ can be computed. In fact solving (3.4) for $D_0{}^+$ yields

$$D_0{}^+ = \frac{\dfrac{D^+(D^+ - A^-)}{2} + \sqrt{\left[\dfrac{D^+(D^+ - A^-)}{2}\right]^2 + (D^+)^2 n_i{}^2}}{\sqrt{n_i{}^2 + \dfrac{D^+(D^+ - A^-)}{2} + \sqrt{\left[\dfrac{D^+(D^+ - A^-)}{2}\right]^2 + (D^+)^2 n_i{}^2}}} \tag{4.1}$$

The plan is therefore self-evident. Samples of germanium of known suitable gallium contents A^- are to be saturated with lithium at various temperatures. If a judicious choice of gallium content is made the lithium will not precipitate when the specimen is cooled. Therefore the value of D^+ characteristic of the saturation temperature can be determined through resistivity measurements performed at room temperature. Taking n_i from Fig. 2 it then becomes possible to calculate $D_0{}^+$ using (4.1).

The crystal specimens employed were cut in the form of small rectangular wafers of dimensions, approximately 1 cm × 0.4 cm × 0.1 cm. On the surfaces of these, small filings of lithium were distributed densely enough so that their average separation was less than the half thickness of the specimen's smallest dimension. The filings were alloyed to the germanium specimen by heating in dry helium for 30 seconds at 530°C. Then the crystals were permitted to saturate with lithium by diffusion from the alloy at some chosen lower temperature. After the period of saturation which ranged from one half hour to as long as 168 days, de-

TABLE II

T °C.	ρ_0 ohm cm	A^- (cm^{-3})	ρ ohm (cm)	D^+ cm^{-3}	D_0^+ (cm^{-3})
25					6.6×10^{13}
100	0.0523	2.2×10^{17}	0.0735	$.9 \times 10^{16}$	2.5×10^{14}
200	0.44	1.3×10^{16}	0.90	7.8×10^{15}	4.6×10^{15}
250	0.1494	4.7×10^{16}	0 652	3.9×10^{16}	2.6×10^{16}
300	0.042	2.9×10^{17}	0.108	2.15×10^{17}	7.3×10^{16}
500	0.00614	4.5×10^{18}	0.0340	4.13×10^{18}	$1\ 7 \times 10^{18}$
608	0.00577	5.0×10^{18}	0.049	4.78×10^{18}	2.8×10^{18}
650	0.00584	4.3×10^{18}	0.0178	3.75×10^{18}	2.4×10^{18}

pending on the temperature, the specimen surface was lapped smooth with carborundum paper. Resistivities were then measured by means of a two point probe.

Table II collects the data showing T, the temperature of saturation in degrees centigrade, ρ_0 the resistivity before saturation, A^- the gallium concentration computed from ρ_0, ρ the resistivity after saturation, and D^+ the lithium concentration computed from ρ. The final column shows D_0^+ computed using (4.1) and Fig. 2.

In Table II the 25°C value of D_0^+ has been taken as the value computed in section III in connection with Fig. 4. It might be thought (in view of a later section in this paper) that the 25° and 100°C values of D_0^+ are not as reliable as the others because at the low temperatures involved the solubility of lithium may be influenced by ion pairing as well as electron-hole equilibria. However, Appendix A shows that the possible error is small.

In Fig. 5 D_0^+ is plotted against temperature using these data. The plot is the curve labeled $Ga^- = 0$, and the open circles were obtained by inserting the measured D^+ values (crosses) into (4.1). We notice that the curve has a maximum in the neighborhood of 600°C. The occurrence of a maximum, is a necessity if D_0^+ is to pass to zero, as it must at the melting point of germanium. It is also worth noticing that D_0^+ near room temperature lies in the range of order 10^{13} cm^{-3}, but that its measurement bas been effected at concentrations as high as 10^{18} cm^{-3}. This illustrates another application of the electron-hole equilibrium, namely in the determination of solubilities.

With D_0^+ in our possession it is interesting to return to (3.4) and to calculate D^+ as a function of temperature for various levels of A^-. This has been done for values of A^- equal to 10^{15}, 10^{16}, 10^{17}, and 10^{18} cm^{-3}. The curves so obtained appear in Fig. 5, labeled $Ga^- = 10^{15}$, 10^{16}, 10^{17}, 10^{18} cm^{-3}, respectively. Their most striking common feature is the minimum which appears below 200°C. This minimum introduces a new prob-

lem in preparing samples without precipitate. Thus consider the $A^- = 10^{16}$ cm^{-3} curve. Suppose the specimen is saturated at 200°C. Then according to Fig. 5, if A^- for the specimen is 10^{16} cm^{-3}, D^+ after saturation will be 7×10^{15} cm^{-3}. However, as the sample is cooled it will tend, at first, to become supersaturated. For example it will achieve its maximum supersaturation at about 140°C. where the minimum of the 10^{16} cm^{-3} curve appears. Thereafter it will return to its undersaturated state. In fact at 25°C a concentration of 9.3×10^{15} cm^{-3} could be supported, whereas the solution contains no more than 7×10^{15} cm^{-3} lithium atoms. Some of these may have precipitated as the cooling process passed through the minimum, so that sufficient time must be provided for the process of re-solution.

If the original saturation had taken place at 250°C, the concentration

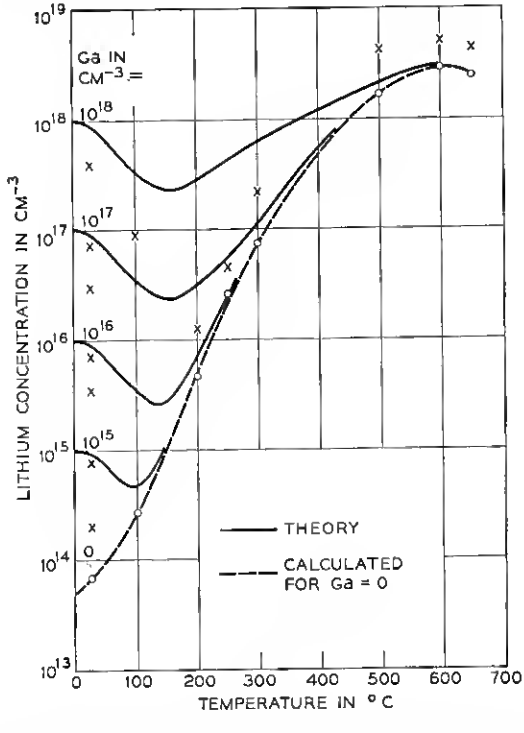

Fig. 5 — Solubility of lithium in germanium as a function of temperature for various gallium dopings. The external phase is an alloy of lithium and germanium. The broken line is the locus of the points (circles) calculated from equation (4.1) for zero gallium concentration. The values of A^- and D^+, used in applying (4.1), correspond to the points shown by × in the illustration. See Table II.

of lithium would have been 2.4×10^{16} cm^{-3}. Since this exceeds the 9.3×10^{15} cm^{-3} supportable at 25°C, such a sample would have contained some precipitate. It was important to avoid these various pitfalls in preparing the specimens used in the above study. Care was taken to insure that this was the case.

We now turn to another application of the electron-hole equilibrium. It has been emphasized that just as a fixed acceptor will increase the solubility of lithium in silicon, a fixed donor should decrease it. In fact in a crystal containing a p-n junction[27] the solubility should be above normal on the p side and below normal on the n side. The built-in field[28] which exists at the junction is a reflection of this difference in solubility, for if it were not present the concentration gradient created by the disparity in solubilities would cause the lithium to diffuse from the p to the n side until its concentration was uniform throughout the crystal. Obviously this field is in such a direction as to cause lithium ions to move back to the p side.*

Now in both silicon and germanium the oxide layers on the surface can react readily with dissolved lithium. As a result the surface behaves as a sink, and at temperatures as low as room temperature lithium is lost to the surface from the body of the crystal. At higher temperatures the body of the crystal can be exhausted of lithium in a few minutes. There are many experiments which one would like to perform in which the crystal must be maintained without loss of lithium at an elevated temperature for long periods of time.

The application now to be discussed involves utilization of the built-in field at a p-n junction to prevent lithium from reaching the surface where

* The distribution of lithium in the space charge region of a p-n junction cannot be computed by the methods advanced thus far. This is because the charge neutrality condition (2.8) is no longer valid. Instead the concentration of lithium is determined by Boltzmann's law,[29] and is given by

$$D^+ = D_*^+ \exp [- qV/kT]$$

where q is the charge on a lithium ion, V is the local electrostatic potential, and D_*^+ is the concentration where V is zero.

V itself must be determined from Poisson's equation[30]

$$\nabla^2 V = -\frac{4\pi\rho}{\kappa}$$

where ρ is the local charge density and κ is the dielectric constant of the medium. In semiconductors ρ is given in terms of V by[31]

$$\rho = q[H + D^+ - 2n_i \sinh (qV/kT)]$$
$$= q[H + D_*^+ \exp [- qV/kT] - 2n_i \sinh (qV/kT)]$$

where H is the local density of fixed donors less the local density of fixed acceptors.

it can attack the oxide. Two specimens of 0.34 obm cm p-type silicon doped witb boron were cut from adjacent parts of a crystal. Each specimen was about 1 cm long, 0.2 cm wide, and 0.15-cm thick. The samples were lapped on No. 400 silicon carbide paper, etched in HF and HNO_3 and sealed in belium-flusbed evacuated quartz tubes, one containing a small grain of P_2O_5 . The tubes were then heated at 1,200°C. for 24 hours. This treatment introduced an n-type layer, highly doped with pbospborus and about 0.001-cm thick, into the surface regions of the specimen in the tube containing P_2O_5 . Upon removal from the tube tbis specimen was lapped on the end to remove the n-skin. Complete removal was determined by testing witb a thermal probe.

Small cubes of lithium (0.038 cm on a side) were placed on the ends of both samples (tbe lapped end of phospborus-doped one) and alloyed to tbe silicon for 30 seconds at 650°C in an atmosphere of dry helium. After this treatment tbe various junction contours sbould have looked like those in Fig. 6, in which the bottom crystal is shown witb the phosphorus-doped skin (cross hatched). During the alloying process a small amount of spherical diffusion of lithium occurs so that small hemispherical n-regions form with the alloy beads as origins. These are shown in Fig. 6.

Next the specimens were heated in vacuum for about six hours at 400°C. Diffusion of litbium into the body of the crystal should occur during this period. However in the sample not protected by the n-type skin lithium should leak to tbe oxide sink on the surface so that the n-type region due to the lithium should bave the pear-shaped contour sbown in the upper part of Fig. 7. If the built-in field at the p-n junction

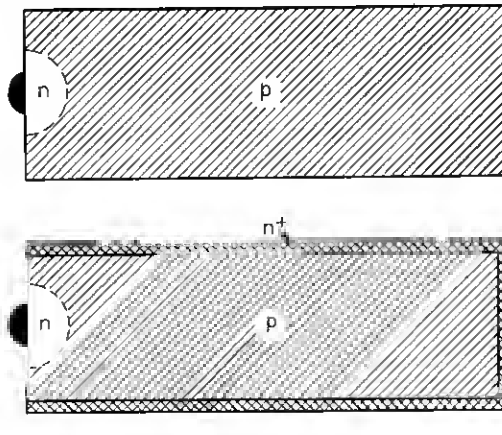

Fig. 6 — Initial stage following alloying in the diffusion experiment to demonstrate the impermeability to lithium of a heavily doped n-type skin on silicon.

formed by the phosphorus layer prevents lithium from reaching the surface, diffusion in the sample with the skin should be plane parallel with a straight front (except at the rear where the skin has been lapped off and lithium can leak out) as the p-n junction contour in the lower part of Fig. 7 indicates.

An acid staining technique[32] which reveals the junction contours should then develop a picture resembling Fig. 7. The two specimens were cut along their long axes and the stain applied to the newly exposed surfaces. The result has been photographed and is shown in Fig. 8 where the crystal on the right has the n-skin. The p-regions show up dark and the n, light. The result agrees with Fig. 7.

In another experiment a crystal completely enclosed in a phosphorus skin was immersed in the tin bath described in Section III. It was discovered that lithium entered the crystal with no evident difficulty, just as though the skin were absent, but once in, could not be driven out by removal of the external source and continued heating. The implication is clear. The built-in field has a rectifying action permitting the lithium to enter the crystal but not to leave. In this sense it performs the same function for the mobile lithium ions as it does for holes in a p-n junction diode.[33]

V. COMPLEX ION FORMATION

In the previous text processes involving the interaction of electrons and holes have been considered. In this section attention will be drawn,

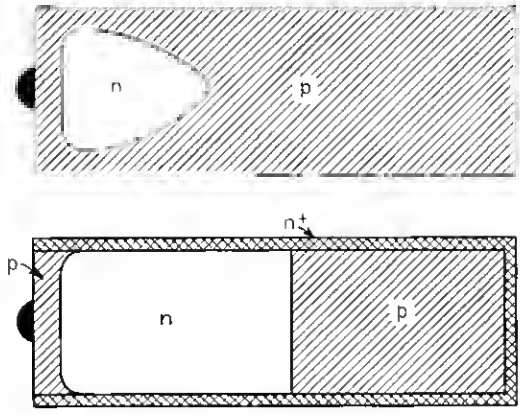

Fig. 7 — Distribution of lithium after an extended period of diffusion at a temperature lower than the alloying temperature — showing leakage out of the crystal in the one case (no-skin) and conservation in the other.

Fig. 8 — Photograph of experimental situation described schematically in Fig. 7.

to the possibility of interactions between the donor and acceptor ions themselves. For example, in (2.1) direct interaction of Li^+ and B^- above 600°C may be possible, especially in view of the mobility of Li^+. Such a reaction was indicated in the work of Reiss, Fuller, and Pietruszkiewicz.[34]

Fig. 9 is of assistance in understanding the nature of these observations. In it are shown plots of the solubility of lithium in silicon. In this case the situation is similar to that involved in the germanium curves of Fig. 5 because the external phase is composed of silicon and lithium and is probably of the liquidus composition. It is formed by simply alloying lithium to the silicon surface. In Fig. 9, Curve A, illustrates how solubility depends on temperature when the silicon is undoped. Curve B, unlike A, is not an experimental plot, i.e., it is not supposed to represent the locus of the points through which it seems to pass. Instead it has been calculated from the theory expounded below. The points themselves are experimental and represent solubility measurements on silicon doped with boron to the level 1.9×10^{18} cm^{-3}.

Curve A possesses a maximum (just as the D_0^+ curve of Fig. 5) in the neighborhood of 650°C. A marked disparity is apparent between solubilities in undoped and doped silicon, the solubility in the latter being greater. Below 500°C this disparity is easily understood. It stems from the electron-hole equilibrium considered previously. However the high solubility in doped silicon at high temperatures is not explicable on this basis since the crystal becomes intrinsic, and e^+e^- is mostly dissociated. To account for this phenomenon Reiss, Fuller, and Pietruszkiewicz invoked the idea of interaction between Li^+ and B^-. They presented the following argument.

At low temperatures lithium ions occupy the interstices of the silicon

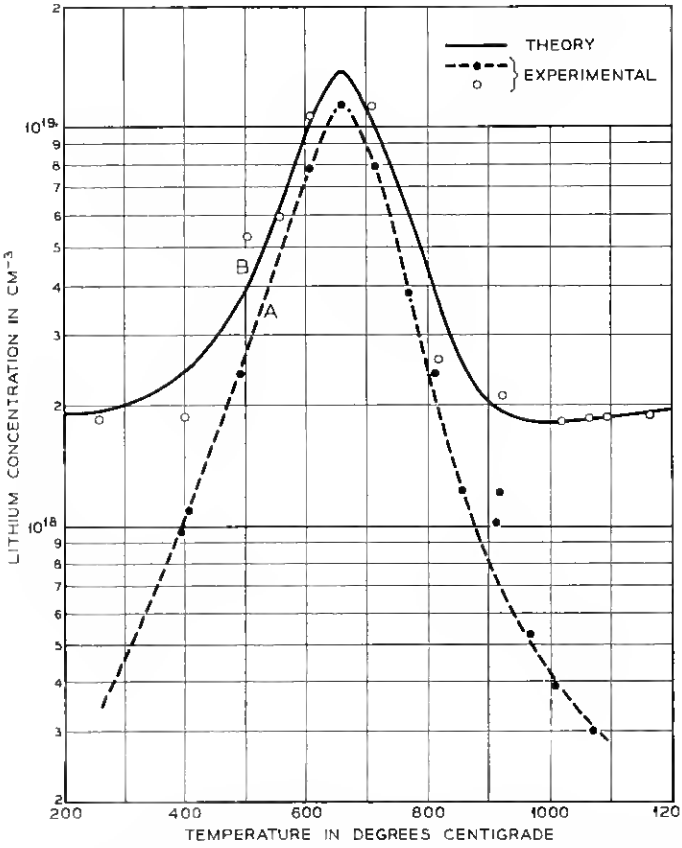

Fig. 9 — Plots showing the solubility of lithium in silicon as a function of temperature. The external phase is an alloy of lithium and silicon. Curve A is for undoped silicon. The locus of the points in B is for silicon doped with about 1.9×10^{18} cm^{-3} boron.

lattice as in Fig. 10. In an interstitial position lithium can approach an oppositely charged boron, but the interaction will be, at the most, coulombic so that an ion pair will form (see later sections). A covalent bond is unable to appear not only because there are no electrons available for it, but also because the lithium ion cannot move to a position where it can satisfy the tetrahedral symmetry inherent in sp^3 hybridization.[35] Calculations (of the sort appearing in the later sections of this paper) show that at high temperatures, at the ion densities involved, ion pairs of the kind depicted in Fig. 10 are completely dissociated.

Suppose, however, that as temperature is raised vacancies dissolve in the silicon lattice, and that one such vacancy occupies a position near

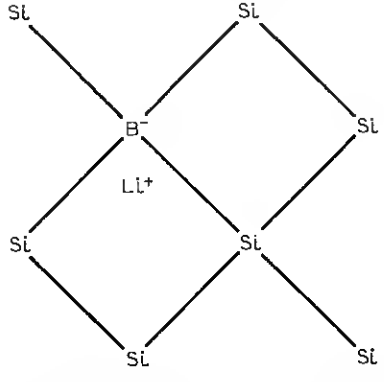

Fig. 10 — Schematic diagram of a silicon lattice showing a lithium ion in an interstitial position near a substitutional boron ion, as it occurs in an ion pair.

a boron ion, as in Fig. 11, a slight modification of Fig. 10 in which the dots represent electrons (dangling bonds). Unpaired electrons such as these might capture an electron from the valence band of silicon so that the vacancy acquires a negative charge and behaves like an acceptor. It is reasonable to suppose that the positive lithium ion will move into this negative vacancy, in the tetrahedral position, and form a covalent bond as in Fig. 11. The lithium-boron complex so formed retains a negative charge and is thus a complex ion. If the specimen were extrinsic at these high temperatures, there would still appear to be as many net acceptors as before the addition of lithium.*

If the LiB^- compound is stable enough (a question to which we shall

* It is possible that rapid cooling may quench some of these LiB acceptors into the crystal at room temperature. If this is so it should be possible to investigate the associated energy level by Hall measurements in the interval of time before the complexes anneal out. Similar phenomena might be observed in germanium.

return below) to hold the lithium atom, the solubility of lithium will be determined principally by the density of boron atoms. At low temperatures, vacancies are reabsorbed and the lithium atoms return to their interstitial positions, at quenched-in densities corresponding to the temperatures of equilibration. However, boron acceptors now appear to be compensated since interstitial lithium behaves as a donor. This renders it feasible to measure the concentration of lithium by the determination of resistivity.

The overall reaction may be written in the form

$$Li^{+} + B^{-} + \square + e^{-} = LiB^{-} \tag{5.1}$$

in which $\square$ represents a vacancy. This equilibrium can be grafted onto (2.1) so that the latter becomes (ignoring un-ionized lithium and boron)

$$\begin{array}{ccccc} Li \text{ (external)} \rightleftarrows & Li^{+} & + & e^{-} & \\ & + & & + & \\ & B^{-} & + & e^{+} & \\ & + & & \uparrow\downarrow & \\ & \square & & e^{+}e^{-} & \\ & + & & & \\ & e^{-} & & & \\ & \uparrow\downarrow & & & \\ & LiB^{-} & & & \end{array} \tag{5.2}$$

The original vertical equilibrium involving holes and electrons loses its significance at high temperatures, and the new vertical reaction becomes important, for both $\square$ and e^{-} appear in increased concentrations. In this way a certain amount of symmetry, insofar as temperature is concerned, is introduced into the problem, i.e., as one equilibrium ceases to dominate

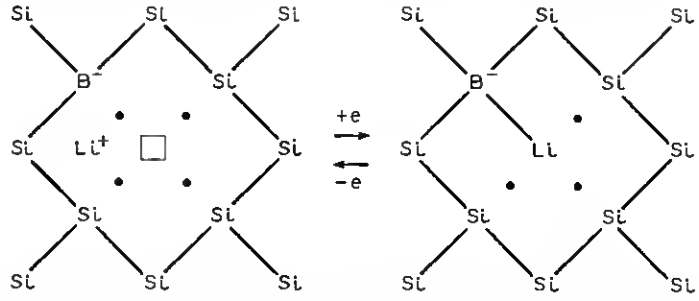

Fig. 11 — Schematic diagram illustrating the reaction in (5.1). The square represents the center of a vacancy and the dots, electrons left unpaired by the occurrence of the vacancy.

the system the other begins to take effect. This symmetry, of course, is necessary for explaining the symmetrical locus of the points around Curve B in Fig. 9.

The scheme (5.2) can be treated quantitatively by applying the mass action principle, but now the symbol D^+ can not be used for the solubility of lithium since the totality of dissolved lithium is distributed between LiB^- and Li^+, and the symbol only applies to the latter. We therefore denote the total concentration of lithium by N_D, and the concentration of LiB^- by C. Then

$$N_D = D^+ + C \tag{5.3}$$

The same argument applies to boron, so that its total concentration will be designated by

$$N_A = A^- + C \tag{5.4}$$

The problem then reduces to specifying N_D as a function of N_A. To accomplish this, to (3.1) and (3.2) is added the mass action expression going with (5.1)

$$\frac{C}{D^+A^-n} = \gamma e^{-(\beta/T)} = \pi \tag{5.5}$$

where γ and β are constants. It has been assumed that the vacancy concentration follows a temperature law of the form $\gamma^* \exp[-\beta^*/T]$ where γ^* and β^* like γ and β are constants. This permits the equilibrium constant when multiplied by the vacancy concentration to assume the form $\gamma \exp[-\beta/T]$ shown in (5.5). In place of (2.8) a new conservation condition,

$$D^+ + p = C + A^- + n \tag{5.6}$$

is introduced. The combination (3.1), (3.2), (5.3), (5.4), (5.5) and (5.6) can be solved so that N_D, the lithium solubility appears as a function of the total boron concentration N_A. Thus

$$N_D = \frac{N_A}{1+\sqrt{1+(2n_i/N_D{}^0)^2}} + \sqrt{\left\{\frac{N_A}{1+\sqrt{1+(2n_i/N_D{}^0)^2}}\right\}^2 + (N_D{}^0)^2} + \frac{\pi N_A (N_D{}^0)^2[1+\sqrt{1+(2n_i/N_D{}^0)^2}]}{2+\pi(N_D{}^0)^2[1+\sqrt{1+(2n_i/N_D{}^0)^2}]} \tag{5.7}$$

In this equation $N_D{}^0$ like D_0 in (3.4) is the solubility of lithium in undoped silicon, i.e., in silicon from which boron is absent.

All the parameters in (5.7) are independently measurable save π

which can be known for all temperatures when γ and β have been determined. Reiss, Fuller, and Pietruszkiewicz used two of the points near Curve B in Fig. 9, above 1,000°C, to define values of N_D for use in (5.7). Then π was computed from (5.7) at these two temperatures. From these values of π, γ and β were determined, and from these, in turn, π was calculated for all temperatures down to 200°C. Using π, N_D was computed from (5.7) over the entire experimental range of temperature. The result is Curve B of Fig. 9 which fits the experimental points very well.

Another check on the validity of the theory (which has not yet been accomplished) would be the following. At high temperatures (5.7) reduces to

$$N_D = N_D^0 + \left\{ \frac{\pi (N_D^0)^2 [1 + \sqrt{1 + (2n_i/N_D^0)^2}]}{2 + \pi (N_D^0)^2 [1 + \sqrt{1 + (2n_i/N_D^0)^2}]} \right\} N_A \qquad (5.8)$$

i.e., N_D is a linear function of N_A with the slope (in brackets) depending upon π. Measurement of this slope at one temperature would thus provide an independent evaluation of π.

A little thought concerning the scheme outlined in (5.2) leads one to wonder why the introduction of boron really increases the solubility of lithium because the same mechanism could be applied to the case in which boron is absent, i.e., to Curve A of Fig. 9. Thus, if B^- is replaced by a silicon atom in Figs. 10 and 11, the entire scheme can be adopted unchanged, except that Si replaces B^-. Thus

$$\begin{array}{lcl} Li \text{ (external)} \rightleftarrows Li^+ & + & e^- \\ + & & + \\ Si & + & e^+ \\ + & & \Updownarrow \\ \square & & e^+e^- \\ + & & \\ e^- & & \\ \Updownarrow & & \\ LiSi & & \end{array} \qquad (5.9)$$

and one wonders why LiB^- should be more stable than $LiSi$. A possible answer is the following:

The tetrahedral covalent radius of boron is 0.88 Å.[36] This is to be contrasted with the tetrahedral radius of silicon which is 1.17 Å.[36] When boron is substituted in the silicon lattice it therefore produces considerable local compressive strain. This strain is partially relieved when a vacancy is formed adjacent to the boron. Thus the energy required to form a vacancy near a boron ion in silicon is less than is required for its

formation near a silicon atom. Hence the endothermal heat of formation of LiB^- in (5.2) is reduced substantially (by the amount of the released energy of elastic strain) below the heat of formation of $LiSi$. This accounts for the greater stability of the former.

The compressive strain around a substitutional boron in germanium is also illustrated by ion pairing studies to be described later in Section XII. Its action in that case keeps the ions which form a pair from approaching each other as closely as they otherwise might. Although really quantitative studies of pairing have not yet been performed in silicon, the lattice parameters of germanium and silicon are sufficiently close to render it fairly certain that the same strain exists in the latter as in the former. This lends support to the previous argument.

Before closing this section there is another related topic which is worth mentioning. This concerns part of the explanation of the retrograde solubility observable in the curves of Figs. 5 and 9, i.e., the occurrence of the maxima. The solubilities along these curves are given by (3.3) in the form

$$D_0^+ = K^*/(K^* + n_i^2)^{1/2}$$

Suppose that at low temperatures K^* is an increasing function of temperature and considerably larger than n_i. Then we have the approximation

$$D_0^+ = (K^*)^{1/2} \tag{5.10}$$

in which the solubility D_0^+ must increase with temperature. If n_i increases more rapidly than K^* with temperature, a point will be reached at which n_i^2 in the denominator of the (3.3) in its special form above, exceeds K^* by so much that the latter can be ignored. When this is so another approximation holds,

$$D_0^+ = \frac{K^*}{n_i} \tag{5.11}$$

in which D_0^+ decreases with temperature since n_i increases more rapidly than K^*. Since (5.10) predicts an increase in solubility with temperature at low temperatures and (5.11) a decrease at higher temperatures a maximum occurs somewhere between. The maximum may not be due to this cause alone, however. For example K^* contains the activity, α, in the external phase, and this may vary with temperature in an erratic manner.

In any event the influence of the electron-hole equilibrium on D_0^+ in both silicon and germanium cannot be ignored. The fact that the distribution coefficients of donors and acceptors in silicon are usually some

ten-fold greater than in germanium may be due to the smaller width of the forbidden gap in the latter. This makes for greater values of n_i and according to (3.3) smaller values of D_0^+.

VI. ION PAIRING

The preceding text drew an analogy between semiconductors and aqueous solutions — phenomena such as neutralization, common ion effects, and complex formation have been discussed. Another feature of "wet" chemistry which has appealed to chemists concerns the influence of coulomb forces among ions on the properties of solutions. This subject is of peculiar interest because such forces are well understood, and considerable progress can be made in the quantitative prediction of their effects.

The first really successful theoretical treatment of coulomb forces in solution is the Debye-Hückel theory.[37] This treatment recognizes the long range character of coulomb forces, and endeavors to account for their effects in terms of a communal interaction involving all of the ions in solution. The theory has now been shown to include certain statistical inconsistencies[38] which, however, are of small consequence in dilute solutions where theory and experiment are in excellent agreement.

The central feature of the Debye-Hückel theory is the concept of the ionic atmosphere, i.e., the time average excess concentration of ions of opposite sign which accumulates in the neighborhood of a particular ion. The radius of this atmosphere is measured (order of magnitude-wise) by the now famous Debye length.

$$L = \sqrt{\frac{k\kappa T}{8\pi q^2 N}} \tag{6.1}$$

in which κ is the dielectric constaut of the medium, q is the charge on an ion, and N is the (in this case identical) concentration of both positive and negative ions. As κ decreases or N increases, L becomes smaller so that the atmosphere is more tightly gathered in. As this process continues a stage is reached in which the atmospheres of some of the ions may be best thought of as being fully constituted by a single ion of opposite sign, i.e., an *ion pair* forms. This pair-wise interaction is so intense relative to the communal interaction mentioned above, that insofar as the paired ions are concerned it may be regarded as the only interaction influencing the distribution of the pairs themselves. Unpaired ions may still be treated by the communal Debye-Hückel theory but their concentration must be considered as the true concentration of ions reduced by the

concentration of pairs since the latter possess effectively no fields. In any event when pairing occurs the Debye-Hückel effects are relatively second order, since, even normally, they represent quite small deviations from ideal solution behavior. Under pairing conditions it is desirable, in the first approximation, to focus one's attention on the pairing interaction.

While developing the aqueous solution analogy inherent in our semiconductor model it is natural to inquire whether or not a system like (2.1), in which at least one of the ions can move, will show effects due to coulomb interaction. A preliminary calculation using (6.1) indicates that if coulomb effects are to be observed they are likely to be of the ion pairing variety rather than of the Debye-Hückel type because the dielectric constants of semiconductors are low relative to that of water, e.g., 12 for silicon[39] and 16 for germanium[40] as against 80 for water.[41] The dominance of ion pairing stems, as it will become clear later, from still another feature peculiar to semiconductors. This is the closeness with which two ions of opposite sign can approach one another in semiconductors. In any event experiments are not yet at the stage of sensitivity necessary for the accurate measurement of the small Debye-Hückel effects so that we are virtually compelled to ignore such phenomena.

Fig. 10 is a picture of an ion pair in boron-doped silicon. Corresponding to this process one may sketch in another vertical equilibrium in (2.1) to yield (ignoring un-ionized Li)

$$\begin{array}{ccccc} Li \text{ (external)} \rightleftarrows & Li^+ & + & e^- & \\ & + & & + & \\ & B^- & + & e^+ & \qquad (6.2) \\ & \Updownarrow & & \Updownarrow & \\ & [Li^+B^-] & & e^+e^- & \end{array}$$

where $[Li^+B^-]$ stands for the ion pair in which the individual ions maintain their polar identities and the binding energy is coulombic. The ion pair is a compound in a statistical sense since as will be seen later the distance between the ions of a pair is distributed over a range of values. The interaction between Li^+ and B^- is to be distinguished from that shown in (5.2). The latter occurs at high temperatures whereas the former is presumably limited to low temperatures, below 300°C.

The quantitative aspects of ion pairing were first considered by Bjerrum[42] and later by Fuoss[43] who placed Bjerrum's theory on a somewhat more acceptable basis. Fuoss's theory, however, suffers from some of the same limitations as Bjerrum's. Nevertheless the Bjerrum-Fuoss theory is capable of satisfying experimental data over broad ranges of conditions. In the next section we present a brief resumé of this theory together with relevant criticism and its relation to a more refined theory due to Reiss.

VII. THEORIES OF ION PAIRING

Fuoss begins by considering a solution of dielectric constant κ, containing equal concentrations, N, of ions of opposite sign. When equilibrium has been achieved each negative ion will have another ion (most probably positive) as a nearest neighbor, a distance r away from it. Fuoss discounts the possibility that the nearest neighbor will be another negative ion, and proceeds to calculate what fraction of such nearest neighbors lies in spherical shells of volumes, $4\pi r^2\,dr$, having the negative ions at their origins. If this fraction is denoted by $g(r)\,dr$, it may be evaluated as follows.

In order for the nearest neigbbor to be located in the volume, $4\pi r^2\,dr$, two events must take place simultaneously. First the volume, $4\pi r^3/3$, enclosed by the spherical shell must be devoid of ions, or else the ion in the shell would *not* be the nearest neigbbor. Since $g(x)dx$ is the probability that a nearest neighbor lies in tbe sbell, $4\pi x^2\,dx$, the probability that a nearest neighbor does not lie in this shell is $1 - g(x)dx$. From this it is easily seen that the chance that the volume $4\pi r^3/3$ is empty is

$$E(r) = 1 - \int_a^r g(x)\,dx \tag{7.1}$$

where a is the distance separating the centers of the two ions of opposite sign when they have approached each other as closely as possible.

The second event which must take place is the occupation of the shell $4\pi r^2\,dr$ by *any* positive ion. The chance of this event depends on the time average concentration of positive ions at r. This concentration is bound to exceed the normal concentration N by an amount depending on r, because of the attractive effect of the negative ion at the origin. It may be designated by $c(r)$. The probability in question is tben

$$4\pi r^2 c(r)\,dr \tag{7.2}$$

The chance $g(r)\,dr$ that the nearest neighbor lies in the shell $4\pi r^2\,dr$ is therefore the product of (7.1) by (7.2), i.e., the product of the probabilities of the two events required to occur simultaneously. This leads to the relation

$$g(r) = \left(1 - \int_a^r g(x)\,dx\right) 4\pi r^2 c(r) \tag{7.3}$$

an integral equation whose solution is

$$g(r) = \exp\left[-4\pi \int_a^r x^2 c(x)\,dx\right] 4\pi r^2 c(r) \tag{7.4}$$

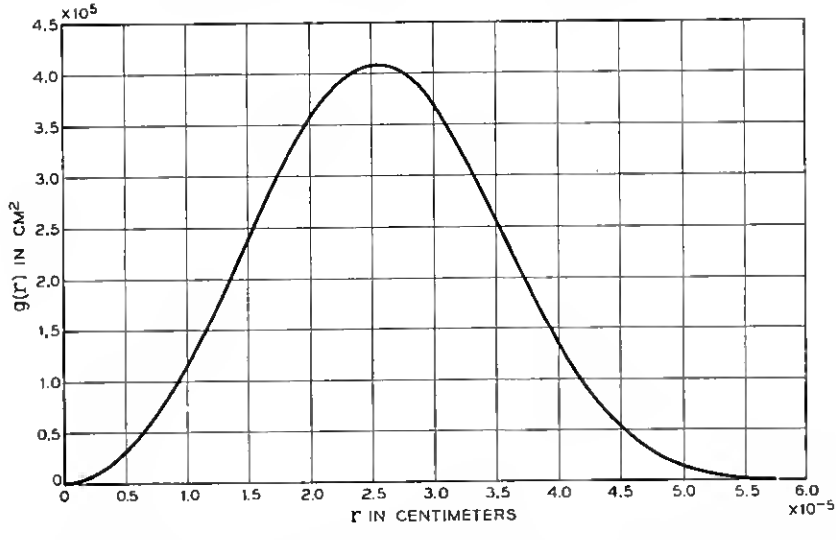

Fig. 12 — Distribution of nearest neighbors in a random assembly of particles for a concentration of 10^{16} cm^{-3}.

That (7.4) solves (7.3) is easily demonstrated by substitution of tbe latter into the former.

If there were no forces of attraction between ions then $c(r)$ would equal N, and if a is take equal to zero (7.4) reduces to

$$g(r) = 4\pi r^2 N \exp(-4\pi r^3 N/3) \tag{7.5}$$

This function is plotted in Fig. 12 for the case $N = 10^{16}$ cm^{-3}. Note that the position of the maximum, tbe most probable distance of location of a nearest neighbor, occurs near tbe value of r equal to $(3/4\pi N)^{1/3}$. This is the radius of the average volume per particle when the concentration is N, i.e. the volume, $1/N$.

In order to write $g(r)$ for the case of coulombic interaction it is necessary to compute $c(r)$ under these conditions. Fuoss (after Bjerrum) reasoned as follows. If a theory can be constructed which depends only upon the characteristics of *near* nearest neighbors (nearest neighbors at small values of r) then tbe force of interaction experienced by the nearest neighbor can be assumed to originate completely in the coulomb field of the negative ion at the origin. This is predicated on the argument that both positive and negative ions develop atmospheres of opposite sign which are superposed when the two ions are close to one another. The result is a cancellation of the net atmosphere leaving nothing for the two

ions to interact with but themselves. Thus the potential energy of interaction, for near nearest neighbors will be

$$-\frac{q^2}{\kappa r} \tag{7.6}$$

For small values of r, therefore, $c(r)$ can be derived from Boltzmann's law and is given by

$$c(r) = h \exp [q^2/\kappa kTr] \tag{7.7}$$

where h is a constant. Guided by the requirement that $c(r)$ should equal N at infinite distance from the central negative ion, h was set equal to N giving, finally,

$$c(r) = N \exp [q^2/\kappa kTr] \tag{7.8}$$

The assumption that a theory could be developed depending only on near nearest neighbors proved reasonable, but the choice of $h = N$ in (7.8) leads to certain logical difficulties. Thus the average volume dominated by a given negative ion is evidently $1/N$. If (7.8) is summed over this volume the result, representing the number of positive ions in $1/N$, should be unity since there are equal numbers of positive and negative ions. Unfortunately, the result exceeds unity by very large amounts except for very small values of N, i.e., for very dilute solutions. We shall return to this point later.

If (7.8) is inserted into (7.4) the resulting $g(r)$ has the form typified by Fig. 13. First, there is an exponential maximum occurring at $r = a$, followed by a long low minimum, and this by another maximum which like the one in Fig. 12 occurs, not far from $r = (3/4\pi N)^{1/3}$, if N is not too large. For small values of N the minimum occurs at

$$r = b = q^2/2\kappa kT \tag{7.9}$$

The function $g(r)$ is actually normalized in (7.4) so that the area under the curve is unity. The second maximum corresponds to the most probable position for a nearest neighbor in a random assembly, i.e., to the maximum in Fig. 12. Essentially the first maximum has been grafted onto Fig. 12 by the interaction at close range which makes it probable that short range neighbors will exist. At high values of N the region under the first maximum becomes so great that enough area is drained (by the condition of normalization) from the second maximum to make it disappear entirely. At this point the minimum is replaced by a point of inflection. More will be said concerning this phenomenon later.

Fuoss chooses to define all sets of nearest neighbors inside the mini-

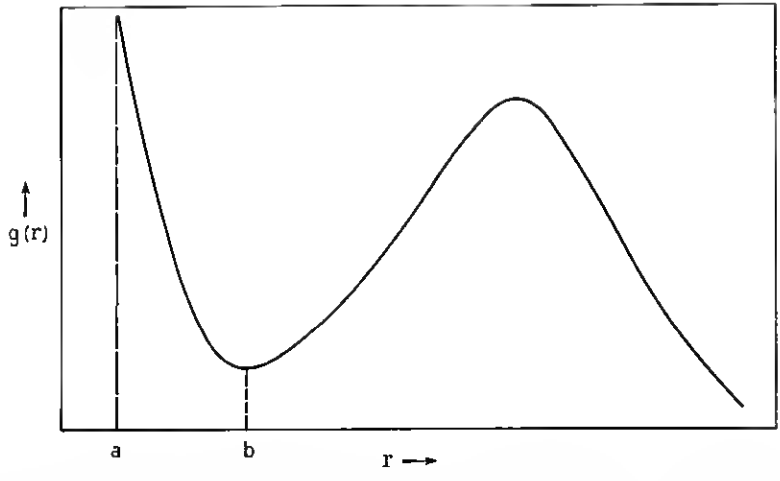

Fig. 13 — Schematic distribution of neighbors in an assembly of particles when forces of interaction are present. Repulsive forces are reflected in the appearance of a distance a, of closest approach of two particles, attractive forces by the exponential maximum at a.

mum, i.e., inside $b = q^2/2\kappa kT$, as ion pairs, and the rest as unpaired. No thought is given to the small fraction of nearest neighbors which involves ions of like sign, as it must be small inside $r = b$. Nor is any thought given to the possibility that a given positive nearest neighbor may be the nearest neighbor of two negative ions simultaneously. Such a coincidence would be very improbable at a distance short enough to be within $r = b$. Thus if the entire theory can be made to depend on what happens inside b, its foundations are reasonable, except for the choice of $h = N$.

To obviate this difficulty Fuoss had further to devise a means of performing all calculations under conditions where the choice of $h = N$ was not inconsistent. He assumed (following Bjerrum) that paired and unpaired ions were in dynamic equilibrium and that the law of mass action could be applied to this equilibrium. Thus if P represents the concentration of pairs, $N - P$ denotes the concentration of unpaired ions of one sign and the mass action expression is

$$\frac{P}{(N - P)^2} = \Omega \tag{7.10}$$

where Ω is an equilibrium constant independent of concentration. At infinite dilution, where the assignment $h = N$ is valid, Ω should be the same as at higher concentrations. Therefore (7.4) can be used to evalu-

ate Ω at infinite dilution, and the value so obtained employed at higher concentrations.

Besides the inconsistency of the choice, $h = N$, the form (7.4) contains another objectionable feature. This is revealed by a more rigorous treatment devised recently by Reiss,[44] and has to do with the factor,

$$\exp\left[-4\pi \int_a^r x^2 c(r)\, dx\right],$$

in (7.4). It can be shown that this factor is inconsistent with the supposition that the nearest neighbor to a given negative ion interacts only with that ion and no other. Fortunately, in Fuoss's scheme $g(r)$ given by (7.4) needs to be used only at infinite dilution, and then only for such values of r as lie inside b. Under this condition and in this range the exponential factor in question can be replaced by unity from which it deviates only slightly. Thus the form of $g(r)$ used eventually is

$$g(r) = 4\pi r^2 N \exp\left[q^2/\kappa kTr\right] \tag{7.11}$$

Ω is computed as follows. At infinite dilution P tends toward zero so that (7.10) becomes

$$\frac{P}{N} = \Omega N \tag{7.12}$$

But P/N is the fraction of ions paired which by definition is the fraction of nearest neighbors lying inside $r = b$. From the definition of $g(r)$, P/N is evidently given by

$$\frac{P}{N} = \int_a^b g(r)\, dr = 4\pi N \int_a^b r^2 \exp\left[q^2/\kappa kTr\right] dr \tag{7.13}$$

which upon substitution in (7.12) yields

$$\Omega = 4\pi \int_a^b r^2 \exp\left[q^2/\kappa kTr\right] dr \tag{7.14}$$

The evaluation of Ω in this way permits one to base the entire theory on the distribution of *near* nearest neighbors, so that all the assumptions which demand this procedure are validated.

Using the computed Ω in (7.10) P can be evaluated, and also $N - P$ which as the concentration of *free* ions of one species measures the thermodynamic activity of that species. In this manner it is possible to calculate the *equilibrium* effects of coulomb interaction insofar as solution properties are concerned. To treat transport phenomena such as ionic mobility in an applied electric field Fuoss assumes that paired ions repre-

senting neutral complexes are unable to respond to the applied field and so do not contribute to the overall mobility. The mobility of unpaired ions is assumed to be μ_0, the mobility observable at infinite dilution. The apparent mobility μ at any finite concentration is then μ_0 reduced by the fraction P/N of ions paired. Thus

$$\mu = [1 - (P/N)]\mu_0 \tag{7.15}$$

The Bjerrum-Fuoss theory when applied to real systems reproduces the experimental data very well, although the parameter a, the distance of closest approach, needs to be determined from the data itself.

The concept of a pair defined in terms of the minimum occurring at b, becomes rather vague when that minimum vanishes in favor of a point of inflection. At this stage triplets and other higher order clusters form and the situation becomes very complicated.

In Reference 44, Reiss has developed a more refined theory of pairing. Instead of avoiding the use of an inconsistent $g(r)$ by introduction of the mass action principle, an attempt is made to provide a rigorous form for $g(r)$, which proves to be the following

$$g(r) = \exp [-4\pi r^3 N/3]\, 4\pi r^2 h \exp [q^2/\kappa kTr] \tag{7.16}$$

in which

$$h = 1 \Big/ \left[\int_a^\infty \exp [-4\pi r^3 N/3]\, 4\pi r^2 \exp [q^2/\kappa kTr]\, dr \right] \tag{7.17}$$

It is also shown that the activity of an ionic species, measured by $N - P$ in the Bjerrum-Fuoss theory, is measured by $\sqrt{hN}$ in the more rigorous theory. The distribution (7.16) suffers neither from an inability to conserve charge in the volume $1/N$ (as does (7.4)) nor from any inconsistency involving the interaction of a nearest neighbor with other ions than the one to which it is nearest neighbor [as does (7.4)].

When $\sqrt{hN}$ computed by (7.17) is compared with $(N - P)$ computed according to (7.10) and (7.14), for arbitrary values of κ, a, T, and N, the results are almost identical. This shows the virtue of the Bjerrum-Fuoss theory, and in fact, suggests that in most cases it should be used for calculation rather than the more refined theory, for the latter involves rather complicated numerical procedures.

The refined theory can also be adapted to the treatment of transport phenomena.[45] Thus in place of $g(r)$ it is possible to write a distribution function $\Gamma(\vec{r})$, specifying the fraction of nearest neighbors lying in the volume element $d\vec{r}$, in a system in the steady state rather than at equilibrium. In the presence of an applied field the distribution loses its spheri-

cal symmetry and it must be defined in terms of the volume elment $d\vec{r}$, lying at the vector distance $\vec{r}$, rather than in terms of the spherical shell of volume, $4\pi r^2\,dr$. In reference (44) it is shown that

$$\Gamma(\vec{r}) = \exp\left[-4\pi r^3 N/3\right] c(\vec{r}) \tag{7.18}$$

where $c(\vec{r})$ is the density function in the non-equilibrium case, and is determined by the equation

$$\frac{kT}{q}\nabla^2 c + c\nabla^2\psi + \nabla c\cdot\nabla\psi = 0 \tag{7.19}$$

after suitable boundary conditions have been appended. The quantity ψ, designates the local electrostatic potential, determined by the ions as well as the applied field. These equations are restricted specifically to the semiconductor case in which the negative ion is unable to move.

The current carried by nearest neighbors in the volume element $d\vec{r}$ in unit volume of solution is

$$J(\vec{r}) = -\exp\left[-4\pi r^3 N/3\right] c(\vec{r})\mu_0\nabla\left[\psi + (kT/q)\,\ell n\, c(\vec{r})\right] \tag{7.20}$$

Using these equations it proves possible in reference 45 to provide a more refined version of (7.15) in which the mobility of nearest neighbors inside $r = b$ need not be considered zero, nor those outside $r = b$ be considered perfectly free and possessed of the mobility μ_0. In fact the average mobility of a nearest neighbor separated by a distance r from its immobile partner proves to be

$$\bar{\mu} = \frac{\mu_0}{2(1-F)}\left(\left[\frac{\varepsilon^2}{3r^2} + \frac{4\varepsilon}{3r} + 2\right]\exp\left(-\varepsilon/r\right) + 2F\left(\frac{\varepsilon}{3r} - 1\right)\right) \tag{7.21}$$

where

$$\varepsilon = q^2/\kappa kT \tag{7.22}$$

and

$$F = \left(\frac{\varepsilon^2}{2a^2} + \frac{\varepsilon}{a} + 1\right)\exp\left(-\varepsilon/a\right) \tag{7.23}$$

For values of r greater than ε (7.21) can be approximated by

$$\frac{\bar{\mu}}{\mu_0} = \frac{1}{2}\left(\frac{\varepsilon^2}{3r^2} + \frac{4\varepsilon}{r} + 2\right)\exp\left(-\varepsilon/r\right) \tag{7.24}$$

and is therefore a function of ε/r. Fuoss's b corresponds to $r = \varepsilon/2$ or to $\varepsilon/r = 2$. Fig. 14 contains a plot of $\bar{\mu}/\mu_0$ versus r for $T = 400°\text{K}$, $a = 2.5 \times 10^{-8}$ cm, $q = 4.77 \times 10^{-10}$ statcoulombs, and $\kappa = 16$. Note that

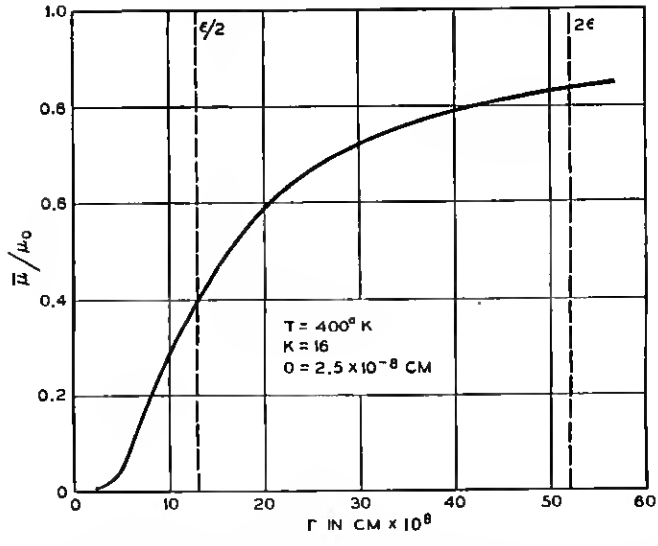

Fig. 14 — Average mobility (calculated from the refined theory of pairing) of a mobile ion in a pair as a function of the distance from its immobile neighbor. The example shown corresponds to a substance having $a = 2.5 \times 10^{-8}$ cm $\kappa = 16$ at a temperature of 400°K.

at $r = \varepsilon/2 = b$, $\bar{\mu}/\mu_0$ is near 0.5 which is the average value of Fuoss's $\bar{\mu}/\mu_0$ for ions taken from either side of $r = b$. Therefore a certain symmetry with respect to $r = b$ does exist, tending to justify Fuoss's model. According to (7.24) $\bar{\mu}/\mu_0$ is 0.8 by the time $r = 3\varepsilon/2 = 3b$, independent of the value of a. In other words an ion located a short distance beyond b *does* have practically complete mobility as the Bjerrum-Fuoss theory assumes.

The refinement of (7.15) which occurs can be written as follows

$$\mu = \left\{ \frac{2h}{(1-F)} \int_a^\infty \left[\left(2r^2 + \frac{4\varepsilon r}{3} + \frac{\varepsilon^2}{3} \right) + 2F \left(\frac{\varepsilon r}{3} - r^2 \right) \exp(\varepsilon/r) \right] \exp(-4\pi r^3 N/3)\, dr \right\} \mu_0 \tag{7.25}$$

Comparison of μ/μ_0 computed from (7.25) with $1 - (P/N)$ appearing in (7.15) over wide ranges of conditions again reveals an excellent correspondence and further substantiates the Bjerrum-Fuoss theory. Since calculations employing the latter are so much simpler it is expedient to regard the cruder theory as an accurate approximation to the more refined one. This practice will be followed from now on.

VIII. PHENOMENA ASSOCIATED WITH ION PAIRING IN SEMICONDUCTORS

In this section we shall discuss some of the phenomena which are to be expected in semiconductors when ion pairing takes place. At the time of writing several of these phenomena have been investigated quantitatively in germanium and casually in silicon. A report on these studies will be given in the later sections of this paper.

In the meantime it is fitting to inquire into the peculiarities which arise because a semiconducting medium rather than a dielectric liquid is involved. The possible means of detecting and measuring ion pairing in semiconductors are numerous, and many of them do not have counterparts in aqueous solution. This implies that a host of new phenomena are to be expeeted, many of which are peculiar to semiconductors.

Some distinctions between semiconductors and liquids are apparent at once. Thus ions are not always mobile in semiconductors at temperatures where ion pairing is pronounced. Lithium is exceptional in this respect, being mohile in germanium and silicon down to very low temperatures. In fact ion pairing has been observed in germanium containing lithium down to dry ice temperatures, and even helow. Another difference is the low dielectric coustant of semiconductors as compared with water. Furthermore, in semiconductors, charge halance need not be maintained by the ions themselves, but may be effected by the presence of holes or electrons. Although charged the latter entities need not be considered in pairing processes since, as particles, they possess effective radii of the order of their thermal wavelengths which may exceed 20 Angstroms at the temperatures involved. At these distances very little eoulomb binding energy would he available. Under certain rare conditions the screening effect of these mobile carriers may make some contribution. This may be particularly the case when *relaxation processes* (to be discussed later) are carried out in poorly compensated specimens of semiconductor, since such processes involve phenomena hetween ions separated by large distances.

A very obvious distinction is the fact that ions in a semiconductor occupy a lattice, and cannot therefore move through a continuum of positions, as in the case of liquid solutions. Furthermore the lattice may introduce elastic strain energy into the binding energy of a pair. This influence will alter the value of a, the distance of closest approach, when the latter is chosen so as to achieve the best fit hetween theory and experiment. As the extent of pairing is extremely sensitive to the magnitude of a, its measurement provides a useful tool for exploring the state of strain in the neighborhood of an isolated impurity. We shall demonstrate

this application later in connection with the strain in the neighborhood of a substitutional boron in germanium.

Aside from its bearing on the minimum distance a, the existence of the lattice will be ignored in the following considerations.

The values of a, typical of semiconductors, are generally of the order of 2 Angstroms as against 6 to 8 Angstroms for ions in liquids. This results from the fact that liquid ions are generally solvated. The consequence to be expected, and indeed found, is that ion pairing will be far more pronounced in semiconductors than in liquids of comparable dielectric constant.

The fact that ions have limited mobilities in semiconductors can be turned to advantage by choosing a system such as lithium and boron in silicon in which only one species of ion, in the case mentioned, lithium, is mobile. Under these conditions it is possible to obviate the clustering phenomenon, mentioned previously, which appears in liquids at high ion concentrations. Clustering is prevented because the immobile ions are uniformly distributed in a random manner, having been grown into the crystals at high temperature where pairing and related processes are unimportant. The obvious complications attending cluster formation can therefore be avoided.

Of course, mobility, being limited to a single species of ion is also an advantage in the theory of the transport phenomena, in such systems.

It is convenient to list some of the effects due to pairing which are to be expected in semiconductors. We do so in the following compilation.

(A) *Equilibrium Phase Relations*

From (6.2) it is apparent that the pairing equilibrium should affect the solubility of lithium in silicon. The same must be true for germanium doped with an acceptor. Although such effects probably occur, they are accompanied by influences arising from the other possible equilibria. As a result the situation is somewhat complex and it is not easy (see Appendix A) to produce experimental conditions under which pairing will be evident. For this reason quantitative investigations along these lines have not yet been attempted.

(B) *Variation of Energy Levels*

When an ion pair is formed of a donor and acceptor, both the donor and acceptor levels are altered. Thus the proximity of the negative acceptor ion increases the difficulty of return to the donor state for an electron, (i.e. the donor level is raised). Likewise the acceptor level is

lowered. In ion pairs it is in fact to be expected that the donor level will be moved up into the conduction band and the aceptor level down into the valence band.* This change in energy level structure should be apparent in Hall coefficient measurements at low temperature. Experiments of this sort bave been conducted and are reported in this paper. Under certain eonditions this phenomenon may be useful for the elimination of trapping[16] levels from the forbidden gap.

(C) Change of Carrier Mobility

Ion pairs possess dipolar fields, and consequently, scattering cross-sections very much smaller than those of point charges. The addition of lithium to a sample under such conditions tbat more than half the added lithium becomes paired should therefore *increase* rather than *decrease* the mobility of holes. The latter effect is tbe one to be expected in tbe absence of pairing. In other words not only earriers but also the scatterers are removed by eompensating tbe aeceptor with donor. Experiments of this sort have been performed. Tbey are described later in this paper. Since they allow us to measure tbe degree of pairing with good accuracy they bave been very valuable in validating the theory, and also in exploring tbe nature of the potential function in the neighborhood of an isolated acceptor.

(D) Relaxation Times

A semiconductor containing unpaired donors and acceptors at one temperature can be cooled to a lower temperature, and the impurities should then pair. If the temperature is lowered sufficiently, the pairing process will be slow enough to be followed, kinetically, by observing any parameter (such as carrier mobility) sensitive to pairing. Experiments of this sort have been performed and will be described later.

The process of pairing can be characterized by a calculable relaxation time, which depends on the acceptor concentration, the diffusivity of the mobile donor, the dielectric constant, and the charges on the ions among other things. The measured time can therefore be used as a means of determining any one of these parameters.

(E) Diffusion

It is evident tbat pairing should reduce tbe diffusivity of a mobile donor. Studies of diffusion in the presence of an immobile acceptor sbould

* A rough calculation indicates that about 0.5 e.v. would be required to place an additional electron on an ion pair.

therefore reveal the action of pairing. Experiments of this sort have been performed and will also be described in tbis paper.

The reduction in the diffusivity of a donor such as lithium may be desirable in certain places.

(*F*) *Direct Transport*

Diffusion studies suffer from the defect that ion pairing produces a concentration dependent diffusivity. (See Appendix B). For this resaon a very desirable measurement would involve determining the amount of a mobile donor like lithium transported by an electric field through a uniformly saturated specimen of semiconductor. Tbis flux, together with information concerning the level of saturation, should provide a direct measure of the mobility of lithium under homogeneous conditions. Formula (7.15) or its refinement (7.25) could then be applied directly to tbe results.

Tbe above list is by no means complete, for there are still otber techniques available for measurement, for example nuclear and paramagnetie resonance. Enough has been given however to indicate the wide range of pbenomena which ion pairing in solids can affect. In liquids, only A and F are of any consequence. It is important to realize that not only do tbese phenomena serve as tools for tbe study of ion pairing, but tbat ion pairing, when properly understood, can serve as a tool for tbe study of the pbenomena themselves.

IX. PAIRING CALCULATIONS

The evaluation of Ω according to (7.14) presents somewhat of a problem because the integral must be arrived at numeriеally. Fortunately, the literature contains tables[47] of the integral in what amounts to dimensionless form. The transformation

$$\xi = q^2/\kappa kTr \tag{9.1}$$

is introduced and then Ω is shown to be given by

$$\Omega = 4\pi[q^2/\kappa kT]^3\, Q(\alpha) \tag{9.2}$$

wbere

$$\alpha = q^2/\kappa kTa \tag{9.3}$$

and $\log_{10} Q(\alpha)$ is tabulated in Table III.

In a specimen in which the numbers of donors and acceptors are un-

TABLE III

α	$\log_{10} Q(\alpha)$	α	$\log_{10} Q(\alpha)$
2.0	$-\infty$	18.0	2.92
2.5	−0.728	20.0	3.59
3.0	−0.489	25.0	5.35
4.0	−0.260	30.0	7.19
5.0	−0.124	35.0	9.08
6.0	0.016	40.0	11.01
7.0	0.152	45.0	12.99
8.0	0.300	50.0	14.96
9.0	0.470	55.0	16.95
10.0	0.655	60.0	18.98
12.0	1.125	65.0	21.02
14.0	1.680	70.0	23.05
16.0	2.275	75.0	25.01
		80.0	27.15

equal* (7.10) may be written as

$$\frac{P}{(N_A - P)(N_D - P)} = \Omega \tag{9.4}$$

where N_A and N_D are, respectively, the total densities of acceptors and donors.

This equation has the following solution for P/N_D, the fraction of donors paired.

$$\frac{P}{N_D} = \frac{1}{2}\left(1 + \frac{1}{\Omega N_D} + \frac{N_A}{N_D}\right) - \sqrt{\frac{1}{4}\left(1 + \frac{1}{\Omega N_D} + \frac{N_A}{N_D}\right)^2 - \frac{N_A}{N_D}} \tag{9.5}$$

Inspection of (9.5) reveals that for given N_A and Ω, P/N_D is a decreasing function of increasing N_D.

Very often, P/N_D is measured in an experiment, and from this it is desired to calculate a, the distance of closest approach. For such purposes the form (9.5) is not very convenient. In fact an entirely different procedure is to be preferred. Suppose P/N_D is denoted by θ, and θ is substituted into (9.4), into which (9.2) has been inserted. We obtain

$$\log_{10} Q(\alpha) = \log_{10}\left[\frac{1}{4\pi}\left(\frac{\kappa kT}{q^2}\right)^3 \frac{\theta}{(N_A - \theta N_D)(1 - \theta)}\right] \tag{9.6}$$

A knowledge of θ thus suffices to determine $\log_{10} Q(\alpha)$, from which, in turn, α can be determined by interpolation in Table III. Then (9.3) can be used for the evaluation of a.

* This is a situation which cannot arise in liquids, since there, charge balance must be maintained by the ions themselves. It can occur when the ions are of different charge, but then things are complicated by the formation of triplets, etc., in addition to pairs.

TABLE IV

T°K	Ω (cm^3)	T°K	Ω (cm^3)
100	2.2×10^{2}	400	2.3×10^{-17}
150	6.45×10^{-7}	500	1.54×10^{-18}
200	3.42×10^{-11}	600	3.0×10^{-19}
225	1.28×10^{-12}	700	1.03×10^{-19}
250	8.79×10^{-14}	800	4.7×10^{-20}
300	1.61×10^{-15}		

Experiments which will be described later indicate that in germanium, gallium and lithium can approach as close as 1.7×10^{-8} cm. Using this value of a, and $\kappa = 16$, $q = 4.77 \times 10^{-10}$ statcoulombs, the values of Ω appearing in Table IV were computed from (9.2)

With these values, P/N_D, the fraction of donors paired can be computed from (9.5) as a function of temperature and N_A for the simplest case, i.e., the one for which $N_A = N_D$. Fig. 15 contains plots showing these dependences. It must be remembered that all other things remaining the same P/N_D will be greater than the values shown in Fig. 15 when $N_D < N_A$.

A rather important integral to which reference shall he made later is

$$I(r_2, r_1) = \int_{r_1}^{r_2} x^2 \exp\,(q^2/\kappa kTx)\, dx \tag{9.7}$$

The integral appearing in (7.14) is a special case of (9.7) with $r_1 = a$, and $r_2 = b$. $I(r_2, r_1)$ has been evaluated over a considerable range. To facilitate matters the transformation

$$x = (q^2/\kappa kT)\,\lambda \tag{9.8}$$

has been employed. In this notation r_1 and r_2 transform to ρ_1 and ρ_2, and

$$I(r_2, r_1) = (q^2/\kappa kT)^3 \int_{\rho_1}^{\rho_2} \lambda^2 \exp\,(1/\lambda)\, d\lambda = (q^2/\kappa kT)^3 i(\rho_2, \rho_1) \tag{9.9}$$

Figs. 16 and 17 contain plots of $i(\rho_2, 0.05)$ out to $\rho_2 = 5$. The choice of ρ_1 equal to 0.05 was rather unfortunate since for $\kappa = 16$, and $T = 300$°K it corresponds to $\rho_1 = 2.5 \times 10^{-8}$ cm. Since acceptors like gallium possess values in respect to lithium as low as 1.7×10^{-8} cm $i(\rho_2, 0.05)$ is not much use in these cases. The choice 0.05 was made before the experimental data on gallium was available. Below we shall describe a method for extending $i(\rho_2, \rho_1)$ to cases where r_1 is less than 2.5×10^{-8} cm.

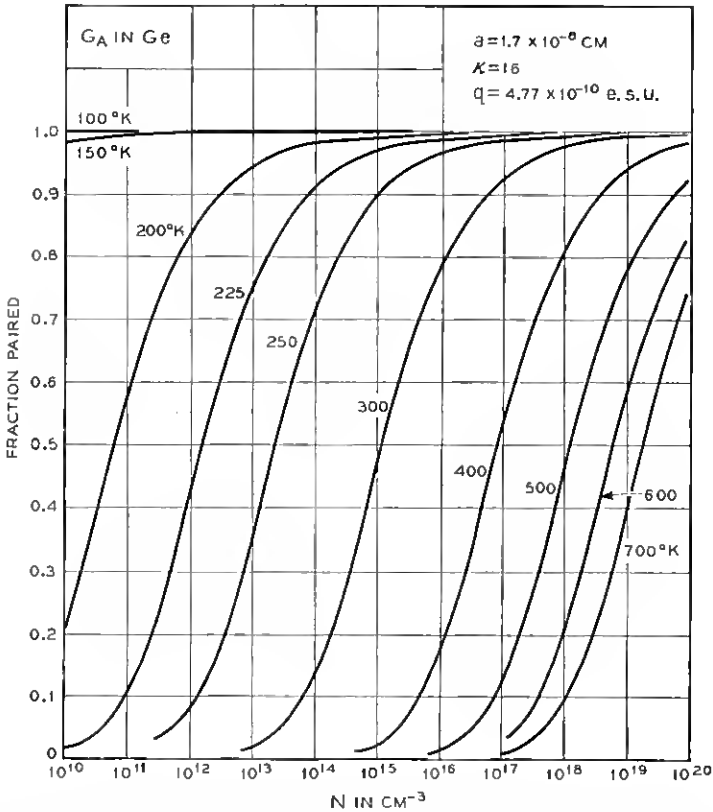

Fig. 15 — Fraction of ions paired, assuming equal densities of positive and negative ions, calculated as a function of temperature and concentration from equation (9.5). The situation illustrated might apply to gallium and lithium in germanium in view of the choice of a and κ.

Fig. 16 covers the range from $\rho_2 = 0.05$ to 0.08 and involves a logarithmic scale because of the sharp variation of i in this range. (This points up the sensitivity of the degree of pairing to the magnitude of a.) Fig. 17 extends the curve to $\rho_2 = 5$. When ρ_2 exceeds 5, $i(\rho_2, 0.05)$ can be obtained from the formula

$$i(\rho_2, 0.05) = 3865 + \frac{(\rho_2)^2}{2} + \frac{(\rho_3)^3}{3} \tag{9.10}$$

In order to determine $i(\rho_2, \rho_1)$ when $\rho_1 \geqq 0.05$, the following formula may be used.

$$i(\rho_2, \rho_1) = i(\rho_2, 0.05) - i(\rho_1, 0.05) \tag{9.11}$$

Finally for cases in which $\rho_1 < 0.05$, Table III can be used. Thus

$$i(\rho_2, \rho_1) = Q(1/\rho_1) - Q(20) + i(\rho_2, 0.05) \tag{9.12}$$

where $1/\rho_1$, and 20 are α values in Table III.

X. THEORY OF RELAXATION

In Section VIII attention was drawn to the fact tbat ion pairing in semiconductors can be made to occur slowly enough so that its kinetics can be followed. It is possible to characterize these kinetics by a relaxation time τ, which we shall endeavor to calculate in the present section.

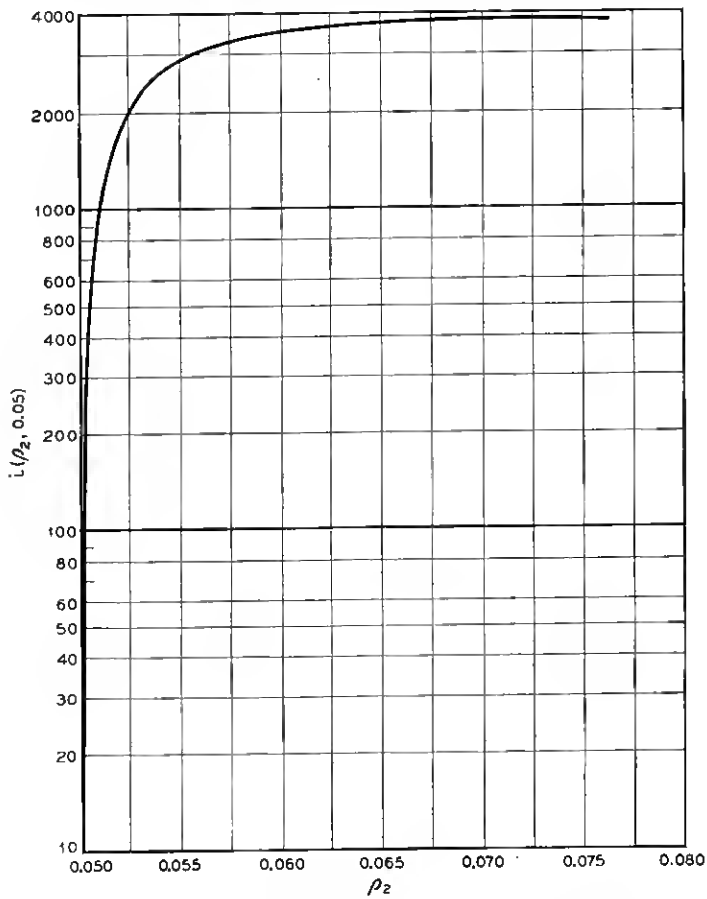

Fig. 16 — Plot, for small values of ρ_2 of $i(\rho_2, 0.05)$ from (9.9).

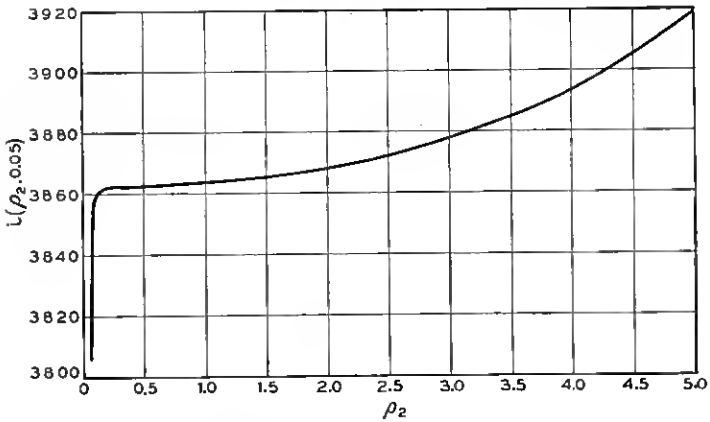

Fig. 17 — Plot, for larger values of ρ_2 , of $i(\rho_2 , 0.05)$ from (9.9).

Suppose a system is first maintained at a temperature high enough to prevent pairing, and then, at an instant designated as zero time, is suddenly chilled to a temperature at which pairing takes place. One thereby has a system which would normally contain pairs but which finds itself with donors and acceptors which are uniformly and randomly distributed. Since the donors are assumed mobile, a process ensues whereby they drift toward acceptors until an equilibrium is established in which each acceptor develops an atmosphere of donors with density $c(r)$, given by (7.7).

This final state in which the atmosphere is fully developed is the paired state characteristic of the lower temperature. The relaxation time to be defined must measure the interval required for the near completion of the above process.

In order to acquire physical feeling for the phenomenon, we begin with some simple considerations. In particular a system will be dealt with containing equal numbers of positive and negative ions. This restriction can be lifted later.

Now, to a first approximation the pairing phenomenon may be regarded as a trapping process in which mobile, positive donor atoms are captured by the negative acceptors. Thus, suppose each acceptor is imagined to possess a sphere of influence of radius R, beyond which its force field may be considered negligible, and inside which a positive ion is to be regarded as captured. This picture immediately emphasizes certain subtleties which require discussion before further progress can be made.

In the crudest sense one might reason that the probability of an encounter between a positive ion and a negative trap would depend on the

product of the densities of both. These densities must be equal because when a positive ion is trapped the resulting ion pair is neutral so that a trap is eliminated simultaneously. If these equal densities are designated by n, we arrive at the second order rate law

$$-\frac{dn}{dt} = k_2 n^2 \tag{10.1}$$

where k_2 is a suitable constant, and t is time.

This law would be perfectly valid if the mean free path of a mobile positive ion were large compared to the distance between ions and the probability of sticking on a first encounter were small. The trapping *cross-section* rather than the movement prior to trapping would determine the trapping rate. In this case the rate would certainly depend on the concentrations of both the traps and the ions being trapped.

On the other hand, in our case, not only is the mean free path of a positive ion much smaller than the distance between ions, but the sticking probability is high. A given ion must *diffuse* or make many random jumps before encountering a trap and upon doing so is immediately captured. Therefore, the rate of reaction is diffusion controlled.

Because of the random jump process a given mobile ion is most likely to be captured by its *nearest neighbor* during the first half of relaxation, and relative to the degree of advancement of the trapping process, the density of traps may be considered constant. This leads to first order kinetics rather than second,* i.e., to

$$-\frac{dn}{dt} = k_1 n \tag{10.2}$$

where n is the density of untrapped ions.

By definition k_1 is the fraction of ions captured in unit time, i.e., the probability that one ion will be captured per unit time. Its reciprocal must be the average lifetime of an ion. This lifetime

$$\tau = \frac{1}{k_1} \tag{10.3}$$

shall be defined as the *relaxation time* for ion pairing. A rough calculation of τ can be made quickly. Thus, suppose that the initial concentrations of donors and acceptors are equally N. About each fixed acceptor can be described a sphere of volume, $1/N$. On the average this sphere should be occupied by one donor which according to what has been said above, will eventually be captured by the acceptor at the center. In the mind, all

* The phenomenon stems from the fact that first and second order processes are almost indistinguishable during the first half of the reaction, but also from the fact that the diffusion control prevents the process from being a true second order one, although its departure from second order may be small.

the spheres can be superposed so that an assembly of donors N in number is contained in the volume $1/N$, at the density N^2. The problem of relaxation is then the problem of diffusion of these donors to the sink of radius R, at the center of the volume. The bounding shell of the sphere may be considered impermeable, thus enforcing the condition that each donor shall be trapped by its nearest neighbor. Since the diffusion problem has spherical symmetry the radius, r, originating at the center of the sink at the origin may be chosen as the position coordinate. At $r = R$, the density, ρ, of diffusant may be considered zero. The radius, L, of the volume, $1/N$, is so large compared to R, that in the initial stages of diffusion L may be regarded as infinite.

In spherical diffusion to a sink from an infinite field, a true steady state is possible, and this steady state is quickly arrived at when the radius, R, of the sink is small.[48] Under this condition concentration is described by

$$\rho = A - \frac{B}{r} \tag{10.4}$$

where A and B are constants. Furthermore at early times n is still N, the initial concentration at $r = L \approx \infty$, so that

$$\rho(\infty) = N^2 \tag{10.5}$$

In addition we know that

$$\rho(R) = 0 \tag{10.6}$$

These boundary conditions suffice to determine A and B in (10.4), and yield

$$\rho = N^2\left[1 - \frac{R}{r}\right] \tag{10.7}$$

Now the rate of capture ($-(dn/dt)$ in (10.2)) is obviously measured by the flux of ions into the spherical shell of area, $4\pi R^2$, which marks the boundary of the sink. This flux is given according to Fick's law[49] by

$$4\pi R^2 D_0 \left(\frac{\partial \rho}{\partial r}\right)_{r=R} = -\frac{dn}{dt} \tag{10.8}$$

where D_0 is the diffusivity of the donor. Substituting (10.7) into (10.8) yields

$$4\pi N^2 R D_0 = -\frac{dn}{dt} \tag{10.9}$$

During the initial stages of trapping the right side of (10.2) may be

written as k_1N, i.e.,

$$k_1N = -\frac{dn}{dt} \tag{10.10}$$

Equating the left sides of (10.9) and (10.10) gives

$$k_1 = 4\pi NRD_0$$

or

$$\tau = \frac{1}{k_1} = \frac{1}{4\pi NRD_0} \tag{10.11}$$

It now remains to choose a value for the capture radius, R. A reasonable guess may be made as follows: Around each acceptor there is a coulomb potential well of depth

$$V = -q^2/\kappa r \tag{10.12}$$

Since the average thermal energy is kT, it seems reasonable to regard an ion as trapped when it falls to a depth kT in this well. Thus, inserting kT on the left of (10.12) and R for r on the right leads to

$$R = q^2/\kappa kT \tag{10.13}$$

and upon substitution in (10.11) we obtain

$$\tau \approx \frac{\kappa kT}{4\pi q^2 ND_0} \tag{10.14}$$

This result, obtained by crude reasoning, is actually quite close to the more rigorous value derived below. Furtbermore, the above derivation is useful in providing insight into the physical meaning of the relaxation time.

Tbe chief difficulty with the preceding lies in the arbitrary choice of R, and is a direct consequence of the long range nature of coulomb forces. Another difficulty arises because the distribution of donors about acceptors is eventually specified by (7.7) so that at $r = R = q^2/\kappa kT$

$$\frac{\partial c}{\partial r} = -\frac{h}{e}\left\{\frac{\kappa kT}{q^2}\right\} \tag{10.15}$$

Since this slope has a negative value tbe trap exhibits some aspects of a source rather than a sink whicb could only produce a positive concentration gradient. This last objection will not be serious when h is very small since, then the final value of $c(r)$ beyond $r = q^2/\kappa kT = R$ will be effectively zero, as would be required for a perfect sink.

The last point raises still another question: What happens when the sink is not perfect, i.e. where the equilibrium state does not involve complete pairing?

All these difficulties can be removed by a more sophisticated treatment of the diffusion problem. Thus, retain the sphere of volume, $1/N$, enclosing N donors at the density N^2. However, the equations of motion of these donors are altered to account for the fact that besides diffusing they drift in the field of the acceptor at the origin. Thus the flux density of donors will be given by

$$\begin{aligned} J^*(r, t) &= -D_0 \left\{ \frac{\partial \rho}{\partial r} + \left\{ \frac{q^2}{\kappa k T r^2} \right\} \rho \right\} \\ &= -D_0 \left\{ \frac{\partial \rho}{\partial r} + \frac{R}{r^2} \rho \right\} \end{aligned} \tag{10.16}$$

where R has been substituted for $q^2/\kappa kT$. Equation (10.16) is obtained by adding to the diffusion component,

$$-D_0 \frac{\partial \rho}{\partial r}$$

of the flux density, the drift component,

$$-\frac{\mu_0 q}{\kappa r^2} \rho,$$

where μ_0 is the mobility of a donor ion and $-q/\kappa r^2$ is the field due the acceptor at the origin. The Einstein relation[50]

$$\mu_0 = qD_0/kT \tag{10.17}$$

has also been used to replace μ_0 with D_0.

The spherical shell bounding the volume, $1/N$, of radius

$$L = \left(\frac{3}{4\pi N} \right)^{1/3} \tag{10.18}$$

is regarded as impermeable, so we obtain the boundary condition

$$J^*(L, t) = 0. \tag{10.19}$$

Furthermore an arbitrary inner boundary, $r = R$, is no longer defined but use is made of the real boundary, $r = a$, i.e., the distance of closest approach, at which is applied the condition

$$J^*(a, t) = 0 \tag{10.20}$$

As before, the initial condition may be expressed as

$$\rho = N^2 \qquad t = 0 \qquad a < r < L \tag{10.21}$$

Tbe continuity equation,[51] in spberical coordinates takes tbe form

$$\frac{1}{r^2}\frac{\partial}{\partial r}\{r^2 J^*\} = -\frac{\partial \rho}{\partial t} \tag{10.22}$$

Substitution of (10.16) into (10.22) gives, finally,

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left\{r^2\frac{\partial \rho}{\partial r} + R\rho\right\} = \frac{1}{D_0}\frac{\partial \rho}{\partial t} \tag{10.23}$$

Equations (10.23), (10.21), (10.20) and (10.19) form a set defining a boundary value problem, the solution of which is $\rho(r, t)$, from which, in turn, $J^*(r, t)$ can be computed. It then remains to compute (dn/dt) in (10.2) from J^*. The former is not simply $4\pi R^2 J^*$ (as in (10.8)) because now J^* is not defined unambiguously, being a function of r. $J^*(R, t)$ migbt be employed but then the method is no less arbitrary than the simple one described above.

Fortunately, nature eliminates the dilemma. It is a peculiarity of spherical diffusion, wben the sink radius is much smaller than the radius of the diffusion field, tbat after a brief transient period, $4\pi r^2 J^*(r)$, except near the boundaries of the field, becomes practically independent of r, and depends only on t. This feature is elaborated in Appendix C. Since in our case the radius of the field is of order, L, and the effective radius of the sink is of order, R, and $L \gg R$, it may be expected that tbis pbenomenon will be observed. In fact its existence has been assumed previously in the derivation of (10.4).

Under sucb conditions it does not matter how the radius of the sink is defined so long as $4\pi R^2$ is multiplied by $J^*(R)$ and not the value of J^* at some otber location.

The boundary value problem, (10.23), (10.21), (10.20), (10.19) is solved in Appendix C, and it is shown there that the value of $4\pi r^2 J^*(r)$ obtained after the transient has passed is closely approximated by

$$4\pi r^2 J^*(r) = -\frac{4\pi q^2 N^2 D_0}{\kappa k T} e^{-t/\tau} \tag{10.24}$$

with

$$\tau = \frac{\kappa k T(N - M)}{4\pi q^2 N^2 D_0} \tag{10.25}$$

where

$$M = 1/4\pi \int_a^L r^2 \exp\left[q^2/\kappa k T r\right] dr \tag{10.26}$$

The close connection between M defined by (10.26) and h defined by (7.17) is apparent. Thus in (7.17) when $r = L$, $\exp[-4\pi r^3 N/3]$ is e^{-1}, and for larger values of r this exponential quickly forces the convergence of the integral. Therefore the values of h and M will be almost equal. This is not surprising since they are meant to be the same thing, i.e., the average concentration, $c(\infty)$, of donors at infinite distance in the equilibrium atmosphere of an acceptor. Both quantities are computed so as to conserve charge in this atmosphere.

At large values of N, M proves to be much smaller than N so that (10.25) reduces to (10.14), validating the crude treatment, for τ in (10.24) is obviously the relaxation time. This is easily seen by writing

$$-\frac{dn}{dt} = -4\pi r^2 J^*(r) = \frac{4\pi q^2 N^2 D_0}{\kappa k T} e^{-t/\tau} \tag{10.27}$$

from which one derives by integration

$$n = M + (N - M)e^{-t/\tau} \tag{10.28}$$

According to (10.28) at $t = 0$, $n = N$, the correct initial density for unpaired ions. At $t = \infty$, $n = M$, also the correct density, i.e., the density at large values of r, when equilibrium is achieved. Obviously τ plays the role of the relaxation time, since by differentiation of (10.28)

$$-\frac{d(n - M)}{dt} = \frac{(n - M)}{\tau} \tag{10.29}$$

which is to be compared with (10.2) and (10.3).

Values of M can be computed using formulas (9.10), (9.11), and (9.12) and Figs. 16 and 17 since the integral in (10.26) is one of the i integrals Fig. 18 shows some values of M, computed in this way for the temperatures 206°, 225°, 250°, and 300°K, for a semiconductor where the value of $a = 2.5 \times 10^{-8}$ cm, $\kappa = 16$, and $q = 4.77 \times 10^{-10}$ statcoulombs. The plots are of M versus N. Note that the values of M are generally much less than N, the disparity increasing with lower temperatures and larger N.

It is also possible to calculate τ for the above system in its dependence upon N and T. To do this the value of D_0 must be known as a function of temperature. Fuller and Severiens[52] have measured the diffusivities of lithium in germanium and silicon down to about 500°K. These data plot logarithmically against $1/T$ as excellent straight lines. In Fig. 19, we show an extrapolation of the line for lithium in germanium down to the neighborhood of 200°K. From this figure it is possible to read values of

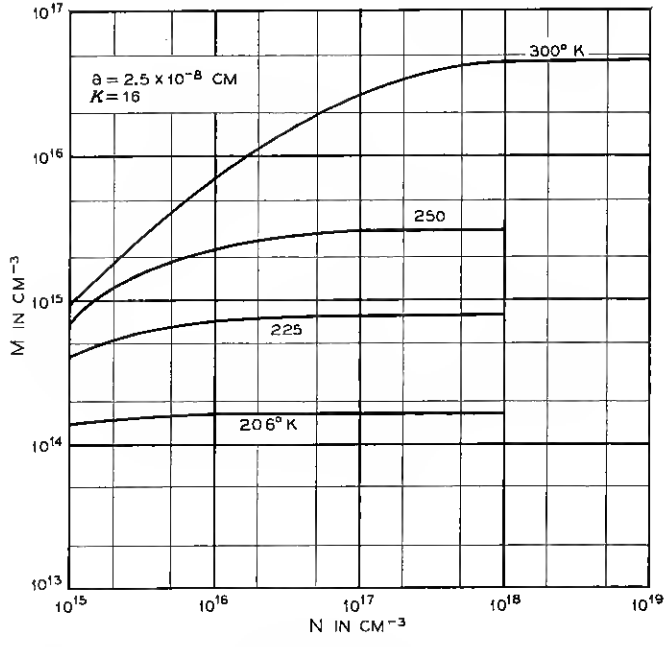

Fig. 18 — Dependence of constant M defined by (10.26) on temperature and concentration, for particular values of a and κ.

D_0 for germanium to which the system of Fig. 18 refers, since κ has been chosen at 16.

Using Figs. 18 and 19, Fig. 20 was computed. It shows τ plotted in seconds versus N for the same temperatures appearing in Fig. 18. These curves show that at values of N as low as 10^{15} cm^{-3} relaxation times are short enough to be observable down to 200°K, being at the most some 50 hours in extent. The value of N makes a big difference. For example at 200°K the relaxation time is only 4 minutes with $N = 10^{18}$ cm^{-3}. Presumably, at 10^{18} cm^{-3}, relaxation could be observed down to much lower temperatures.

It is interesting to note that insofar as M hardly appears in τ, the latter is independent of the distance of closest approach, a. Since a is to some extent empirical this is a fortunate circumstance, and the measurement of τ may provide an accurate means of determining, N, D_0, κ, or q, whichever parameter is regarded as unknown. Furthermore κ as a macroscopic parameter has real meaning in τ since the forces involved may be regarded as being applied over the many lattice parameters separating the drifting donor from its acceptor.

This section will be closed by indicating how the restriction to systems containing equal numbers of donors and acceptors might be lifted. Thus, suppose N_A exceeds N_D. Then there will be $N_A - N_D$ mobile holes maintaining charge neutrality. To a first approximation these will screen the $N_A - N_D$ uncompensated acceptor ions so that the N_D donors will see effectively only N_D acceptors. Thus in first approximation τ can be computed for this system by replacing N in the preceding formulas by N_D.

Of course it is possible that there will be a further effect. Thus the mobile holes will probably shield some of the compensated acceptors as well. This will lead to a further (probably small) reduction in τ, over and above that obtained by replacing N by N_D. We shall not go into this in the present paper, because in most of the experiments performed N_D was near N_A. In the few exceptions the crude correction, suggested above, can be used.

XI. INVESTIGATION OF ION PAIRING BY DIFFUSION

Most of the theoretical tools required for the study of ion pairing have now been provided, and attention will be turned to experiments which

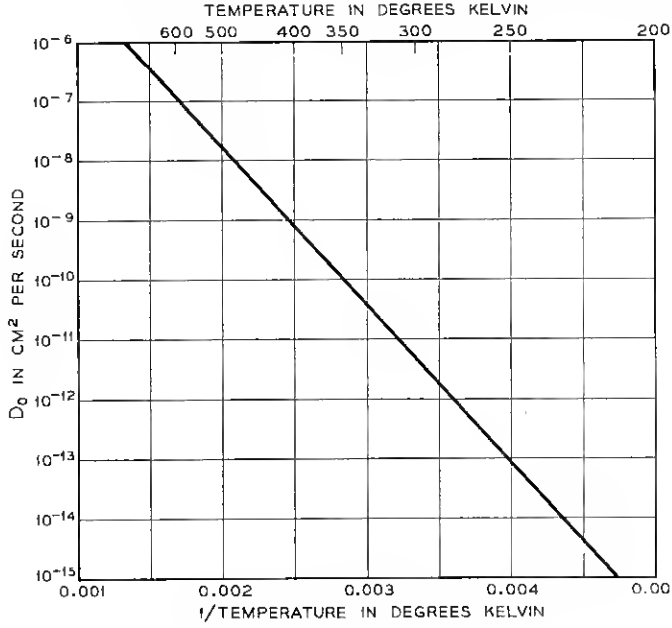

Fig. 19 — Diffusivity of lithium in germanium extrapolated from the data of Fuller and Severiens.

have been performed in this field. A fairly large group of these exist, and it remains to describe them in detail. We shall begin with the study of the diffusion of lithium in p-type germanium.

At the outset a matter having to do with the *diffusion potential* demands attention. This is the potential which arises, for example, in p-type material, because the mobility of a hole is so much greater than the mobility of a lithium ion. In consequence, holes diffuse into regions containing high concentrations of lithium more rapidly than lithium ions can diffuse out to maintain space charge neutrality. As a result such re-

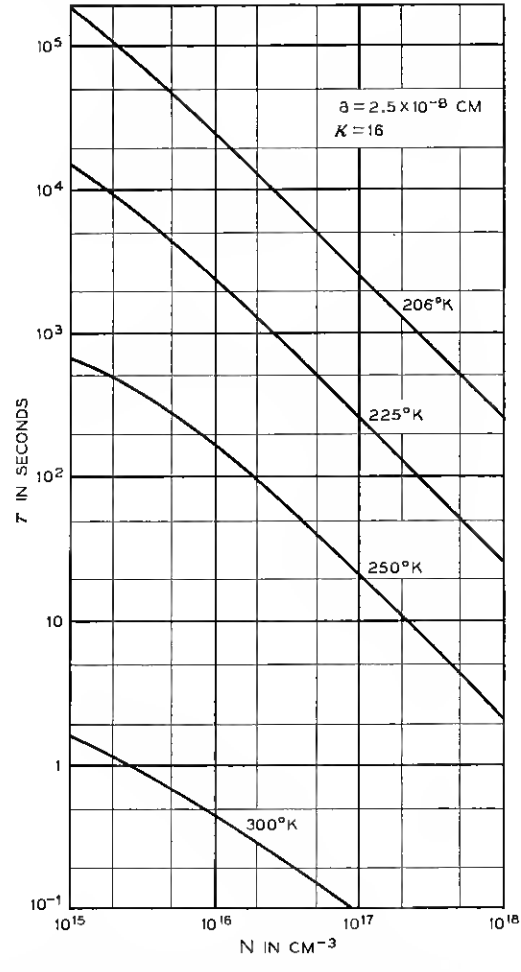

Fig. 20 — Relaxation time as a function of temperature and concentration computed from equation (10.25) using the data of Figs. 18 and 19.

gions develop positive potentials and a field exists tending to expel lithium. This causes the lithium to drift as well as diffuse so that Fick's law[49] is no longer valid.

The most that can be done toward the elimination of diffusion potentials is to minimize them so that no local space charge exists. At equilibrium, this corresponds to the condition[53]

$$N_D - N_A = 2n_i \sinh(qV/kT) \tag{11.1}$$

where V is the local electrostatic potential. It is always permissible to assume that fast moving electrons and holes are in equilibrium relative to diffusing ions. If a material which is p-type everywhere is being considered, (11.1) can be simplified to

$$N_A - N_D = n_i \exp[-qV/kT] \tag{11.2}$$

In Appendix D it is proved that (11.2) will be valid everywhere within a region where N_A is constant and greater than N_D, provided that N_D does not fluctuate through ranges of the order N_A in a distance less than

$$\ell = \sqrt{\frac{\pi\kappa kT}{q^2 N_A}} \tag{11.3}$$

Under most conditions of experiment ℓ will he of the order of 10^{-5} cm. Unfortunately many of the experiments described in this section (particularly those performed at 25°C.) involve diffusion layers as thin as 10^{-6} cm. As a result space charge will exist and the diffusion potential will not always be minimized. Even if it is minimized so that (11.2) is satisfied the residual field will still aid diffusion and lead to higher apparent diffusivities. Therefore the effect cannot be ignored even when minimization has been achieved.

In the absence of space charge the drift component of flux density due to the field is easily computed. It will be given by

$$-\mu \frac{\partial V}{\partial x} N_D \tag{11.4}$$

According to (11.2)

$$-\frac{\partial V}{\partial x} = \frac{kT}{q(N_A - N_D)} \frac{\partial N_D}{\partial x} \tag{11.5}$$

so that (11.4) becomes

$$\begin{aligned} -\frac{\mu kT}{q}\left(\frac{N_D}{N_A - N_D}\right)\frac{\partial N_D}{\partial x} &= -\frac{\mu_0 kT}{q}\left(1 - \frac{P}{N_D}\right)\left(\frac{N_D}{N_A - N_D}\right)\frac{\partial N_D}{\partial x} \\ &= -D_0\left(1 - \frac{P}{N_D}\right)\left(\frac{N_D}{N_A - N_D}\right)\frac{\partial N_D}{\partial x} \end{aligned} \tag{11.6}$$

where (7.15) and the Einstein relation[50] have been used, and D_0 is the diffusivity in the absence of pairing.

P/N_D in (11.6) can be evaluated using (9.5) so that the coefficient preceding $(\partial N_D/\partial x)$ contains N_D as the only variable.

In Appendix B it is shown that ion pairing itself leads to severe departures from Fick's law.[49] In fact the diffusion flux density in the presence of pairing is given by

$$-\frac{D_0}{2}\left(1+\frac{\frac{1}{2}\left(N_D-N_A+\frac{1}{\Omega}\right)}{\sqrt{\frac{1}{4}\left(N_D-N_A-\frac{1}{\Omega}\right)+\frac{N_D}{\Omega}}}\right)\frac{\partial N_D}{\partial x} \tag{11.7}$$

Here again the diffusivity is specified by the factors preceding $(\partial N_D/\partial x)$ and, though variable, depends only on N_D, the local concentration of diffusant. Adding the two coefficients appearing in (11.6) and (11.7) the value of the diffusivity, D, in the presence of both pairing and diffusion potential is obtained. Thus

$$D=\frac{D_0}{2}\left(1+\frac{\frac{1}{2}\left(N_D-N_A+\frac{1}{\Omega}\right)}{\sqrt{\frac{1}{4}\left(N_D-N_A-\frac{1}{\Omega}\right)^2+\frac{N_D}{\Omega}}}+2\left(1-\frac{P}{N_D}\right)\left(\frac{N_D}{N_A-N_D}\right)\right) \tag{11.8}$$

It is obvious from (11.8) that even in the absence of space charge D is an extremely complicated function of N_D, and will be much more complex if space charge needs to be considered. When $N_D \ll N_A$ (11.8) reduces to

$$D=D_0\left(\frac{1}{1+\Omega N_A}\left(1+\frac{N_D}{N_A}\right)\right)\approx\frac{D_0}{1+\Omega N_A} \tag{11.9}$$

Comparison with equation (B15) shows that when (11.8) is true (i.e., in the absence of space charge) the diffusion potential may be ignored for $N_D \ll N_A$. Comparison of (B14) with (B15) shows how much D can vary with N_D when ion pairing occurs.

The proper study of diffusion in the presence of ion pairing should be augmented by a mathematical analysis, accounting for the concentration dependent diffusivity. Since this dependence is complicated the resulting boundary value problem must be solved numerically, and this

represents a formidable task. Although work along these lines is being done we shall content ourselves, in this article, with a less quantitative approach. The following plan has been followed.

A rectangular wafer of semiconductor uniformly doped with acceptor to the level, N_A, is uniformly saturated with lithium to a level, N_D, slightly less than N_A. Thus, the resulting specimen is well compensated but not converted to n-type. Lithium is then allowed to diffuse out of the specimen, and because of the thinness of the wafer, this process may be regarded as plane-parallel diffusion normal to its large surfaces. Low resistivity p-type layers therefore develop near the surfaces. If the thin ends of the wafer are put in contact with a source of current, current will flow parallel to its axis, so that the equipotential surfaces will be planes normal to this axis. The flow of current will be one dimensional because the inhomogeneity in lithium distribution occurs in the direction normal to its flow (see Fig. 21).

If two probe points are placed at a fixed distance apart on the broad surface of the wafer (see Fig. 21), then the conductance measured between them is a reflection of the total number of carriers in the low resistivity layers, i.e., a measure of the total amount of lithium which has diffused out. A more detailed connection between this conductance and diffusivity is derived in Appendix E. For the moment, however, attention will be confined to the description of the general plan of experiment.

According to the formulas derived in the early parts of this section, and also to (B14) and (B15), the diffusivity is something like $D_0/2$ in the

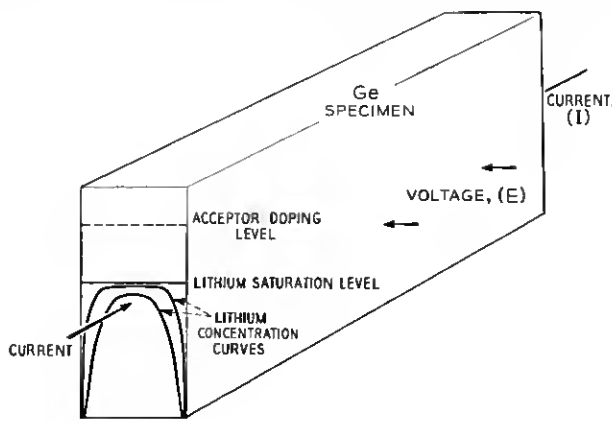

Fig. 21 — Diagram illustrating measurement of dependence of diffusivity on ion pairing (see Section XI).

bulk of the wafer where N_D almost equals N_A, but is as low as $D_0/(1 + \Omega N_A)$ near the surface where $N_D \ll N_A$. If ΩN_A is very much larger than unity as it will be under conditions where appreciable pairing occurs, the diffusivity will, therefore, be much smaller near the surface than at the high end of the diffusion curve, deeper within the specimen. The surface will then offer resistance to diffusion, and it may be expected that the measured value of the diffusivity will correspond more closely to the slow process near the surface rather than to the faster process occurring deeper in the semiconductor. Of course this cannot be entirely true because the resistance at the surface coupled with the lack of resistance inside the wafer will tend to steepen the concentration gradient near the surface. This will give the impression of a diffusivity somewhat higher than the one corresponding to the surface.

If the current flowing in the wafer under the conditions of measurement is I, and the potential measured between the points is V, then the conductance between the points is

$$\Sigma = I/V. \tag{11.10}$$

In Appendix E it is shown (under the assumption that D is constant) that

$$\Sigma/\Sigma_0 = 1 + \frac{2.256\vartheta\sqrt{D}}{d}\left(\frac{\Sigma_\infty N_D^\circ}{\Sigma_0 N_A}\right)\sqrt{t} \tag{11.11}$$

where Σ_0 is the conductance after the specimen is saturated with lithium, but before any lithium has diffused out, and Σ_∞ is the conductance before lithium has been added. N_A is the uniform concentration of acceptor, and N_D° is the initial uniform concentration of lithium, while d is the thickness of the wafer. ϑ is a correction factor which arises because the mobility of holes varies from point to point in the wafer, as the density of lithium varies. There are two extreme types of variation.

The first takes place in a specimen in which, at room temperature (where the conductance measurement is made) ion pairing is complete. Then the local density of impurity scatterers[54] will be $N_A - N_D$. At the other extreme no ion pairing occurs, and the density of scatterers is $N_A + N_D$.

The nature of ϑ depends on how much pairing is involved. In Fig. 22 ϑ has been evaluated in its dependence on N_D° for the extreme cases mentioned. Furthermore it has been assumed then that N_D is given by a Fick's law solution of the diffusion problem, and that diffusion begins in a nearly compensated specimen.

The first thing to notice is that ϑ is not very different from unity in

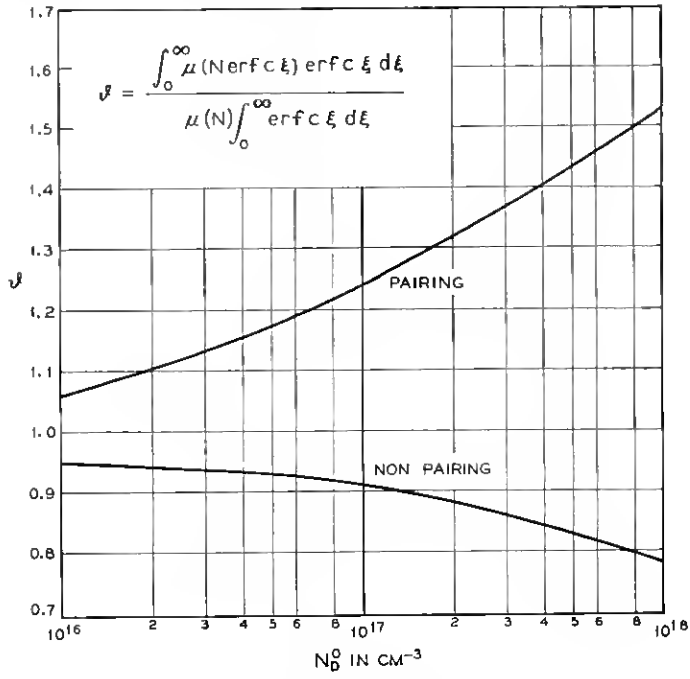

Fig. 22 — Plots of correction factor ϑ, required to compensate for the dependence of hole mobility on the density of scattering centers along a diffusion curve. ϑ is plotted against the initial density of donor and is shown for the two extreme cases of pairing and no pairing.

either extreme, and therefore eloser to unity in some intermediate situation. In any event the correct value of ϑ can be read from Fig. 22 if the experiments involve either extreme at the measurement temperature. This has, in fact, been approximately the case in our experiments, in which pairing is almost complete at the temperature where conductanees have been measured.

According to (11.11) a plot of Σ/Σ_0 against $\sqrt{t}$ should be a straight line of slope

$$S = \frac{2.256\vartheta\sqrt{D}}{d}\left(\frac{\Sigma_\infty N_D^0}{\Sigma_0 N_A}\right) \tag{11.12}$$

Measurement of S therefore affords a measure of D. Of course the apparent D obtained in this manner can never represent anything beyond some average quantity having the general significance of a diffusivity. This follows from the previous diseussion concerning the non-eonstancy

of D. The only exception to this statement occurs in connection with high temperature experiments (above 200°C.) where both pairing and the diffusion potential are of little consequence. The mere fact that Σ/Σ_0 plots as a straight line against $\sqrt{t}$ is not evidence for the constancy of D. In Appendix E it is shown that a straight line will result, even when ion pairing is important, provided that the diffusion potential is based on the no-space-charge condition, i.e. provided that D varies only through its dependence on N_D.

On the other hand, the last statement implies that the existence of a straight line relationship is evidence that the diffusion potential has at least been minimized.

The most careful experiments were performed in germanium doped to various levels with gallium, indium, and zinc as acceptors. The germanium specimens were cut in the form of rectangular wafers of approximate dimensions (1.25 cm × 0.40 cm × 0.15 cm). Fresh lithium filings, were evenly and densely spread on one surface of the wafer, and alloyed to the germanium by heating for 30 seconds at 530°C in an atmosphere of dry flowing helium. Then the other surface was subjected to similar treatment.

After this the specimen was sealed in an evacuated pyrex tube and heated at a predetermined temperature for a predetermined period of time. The temperature was chosen, according to Fig. 5, so that the saturated specimen would still be p-type and just barely short of being fully compensated. Also attention was paid to the problem of avoiding precipitation on cooling. The time of saturation was determined from an extrapolation of the known lithium diffusion data, in germanium, of Fuller and Severiens[52] which is plotted in Figure 19 for the range extending from about 0° to 300°C.

After saturation the sealed tube was dropped into water and cooled. It was opened and the wafer ground on both sides, first with No. 600 Aloxite paper, and then with M 303½ American Optical corundum abrasive paper. The final thickuesses of the specimens ranged from 0.025 to 0.075 cm, the thinnest samples being used for the runs at the lowest temperature.

If the specimen is quite thin and highly compensated it is possible in principle to measure very small diffusivities (as low as 10^{-14} cm^2/sec) within a period of several hours. This is so because the low resistivity layer formed near the surface, although thin, will carry a finite share of the current in thin compensated specimens. On the other hand, additional difficulties arise. Diffusion layers as small as 100Å may be involved. If the surfaee is microscopically rough, diffusion will not be plane-parallel

and the measured diffusivity will appear larger than the real diffusivity. This condition can be partially corrected by etching the surface chemically until it is fairly smooth.

When dealing with such thin layers, the no-space-charge assumption becomes invalid and the diffusion potential ought really to be considered. Considering all the difficulties, i.e., concentration dependence of diffusion coefficient, possible existence of space charge, and roughness of surface, it is apparent that only qualitative effects are to be looked for in the diffusivities which have been measured.

The most that can be predicted is that for specimens containing a given amount of acceptor, the measured D (some average quantity) should be less than D_0, the disparity increasing with decreasing temperature. At high temperatures D should converge on D_0. Furthermore, at a given temperature D should decrease with an increase in concentration of acceptor. These tendencies are in line with the idea that reduction of temperature or increase of doping leads to an increase in pairing.

Runs were carried out on specimens etched with Superoxol[55] at the temperatures 25°, 100°, and 200°C. In the 25°C run the wafer was allowed to remain in the measuring apparatus under the two probe points in air, and Σ was measured from time to time. At 100°C the specimen was immersed in glycerine containing a few drops of HCl, the temperature of the bath being controlled. Periodic removal from the bath facilitated the measurement of Σ. At 200°C glycerine was again used as a sink for lithium, the sample being removed periodically for measurement.

Fig. 23 illustrates some typical plots of Σ/Σ_0 versus $\sqrt{t}$. They are all satisfactorily straight. Fig. 24 shows a plot of log D_0 against $1/T$, extrapolated from the data of Fuller and Severiens.[52] In this illustration, values of log D (obtained from the above measurements by determining the slopes S and employing (11.12)) are also plotted at the temperatures of diffusion. For ϑ the case of complete pairing was assumed.

The first thing to note is that the points for log D all lie below log D_0 except at 200°C and satisfy the qualitative requirement outlined above.* Moreover they drop further below log D_0 as the temperature is reduced, while at 200°C they have almost converged on log D_0.

The results for zinc are particularly interesting. Zinc is supposed to have a double negative charge in germanium.[64] Hence we would expect very intense pairing to occur. This is indicated in the diffusion data where the sample containing zinc at the rather low level, $N_A = 2.7$

* The long range nature of the interaction forces becomes evident when one considers that the diffusivities are being altered by impurity (acceptor) concentrations of the order of 1 part per million.

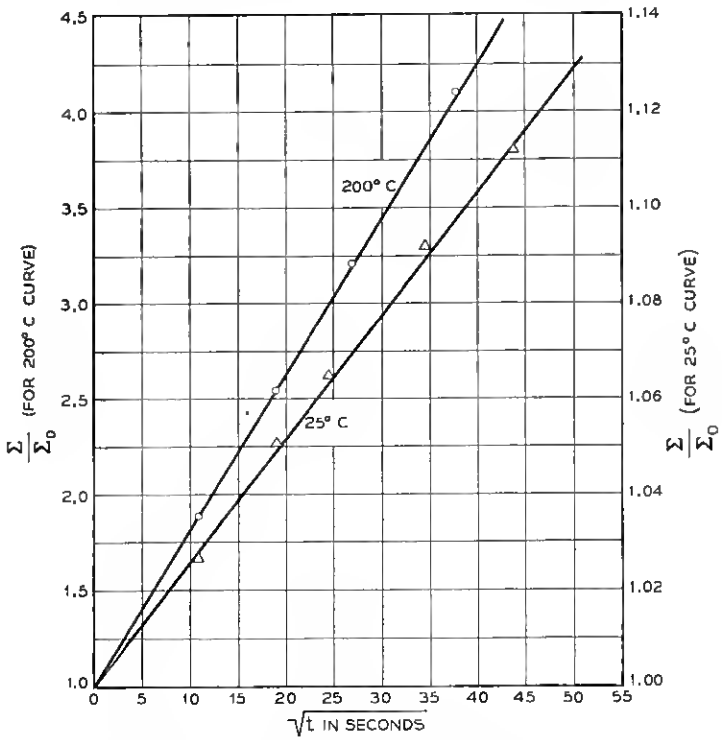

Fig. 23 — Curves illustrating the observed linear dependence of Σ/Σ_0 on the $\sqrt{t}$.

$\times 10^{16}$ cm^{-3}, shows a large reduction in diffusivity even at temperatures as high as 200°C.

The difficulties discussed in this section serve to emphasize the importance of a direct transport experiment in which lithium atoms *uniformly* distributed throughout germanium or silicon, uniformly doped with acceptor, are caused to migrate by an electric field, and their mobilities measured. Because of the uniform dispersion of solutes the mobility will be constant everywhere. Furthermore no diffusion potential will be involved, and also the refined formula (7.25) can be applied. There are, however, many difficulties associated with the performance of this type of measurement.

In closing it may he mentioned that a few much less careful experiments of the kind described here have been performed in boron-doped silicon. The results indicate ion pairing in a qualitative way but more definite experiments are needed.

XII. INVESTIGATION OF ION PAIRING BY ITS EFFECT ON CARRIER MOBILITY

In Section VIII attention was called to the fact that ion pairing should influence the mobility of holes, because each pair formed, reduces the number of charged impurities by two. Thus, a specimen previously doped with acceptor, might, if sufficient lithium is added, exhibit an increase in hole mobility, even though the addition of lithium implies the addition of more impurities. This effect has been observed in connection with the Hall mobility of holes in germanium.

Two specimens of germanium were cut from adjacent positions in a single crystal doped with gallium to the level 3×10^{17} cm^{-3}. One of these was saturated with lithium through application of the same procedure

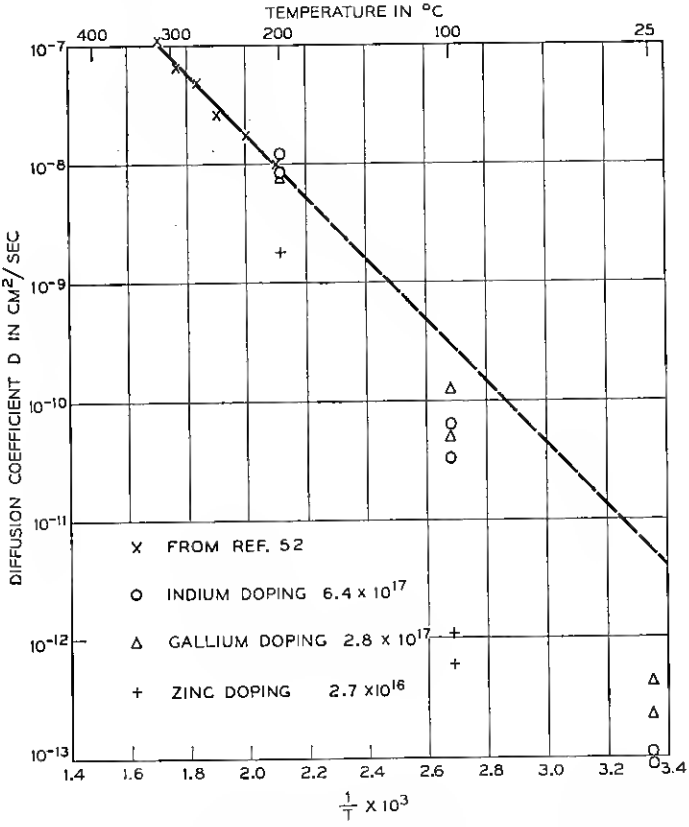

Fig. 24 — Plot of diffusivity of lithium in undoped germanium as a function of temperature — also showing points for apparent diffusivities of lithium in variously doped specimens.

employed in section V. Hall mobilities of the two specimens were measured[65] down to below 10°K. Cooling was carried out slowly to permit as much relaxation into the paired state as possible (see Section X). In Fig. 25 plots of the Hall mobilities versus temperature of both specimens are presented. Curve A is for the sample containing 2.8×10^{17} cm^{-3} lithium. It therefore contained about 5.8×10^{17} cm^{-3} total impurities as compared to the control sample whose curve is shown as B in Figure 25 and which contained only 3×10^{17} cm^{-3} impurities.

The lithium doped bridge exhibits by far the higher Hall mobility for holes (except at very low temperatures where poorly understood phenomena occur). In fact at 40°K the sample containing lithium shows a hole mobility 16 times greater than that of the control at the corresponding temperature. Rough analysis of the relative mobilities at $T = 100$°K indicate $\sim 2 \times 10^{17}$ cm^{-3} scattering centers in the control sample and 5×10^{15} cm^{-3} scattering centers in the sample containing pairs.

This experiment has been repeated with other specimens doped to different levels with gallium and even with other acceptors, and leaves no doubt that a mechanism which is most reasonably assumed to be pairing, is removing charged impurities from the crystal.

The phenomenon we have just described suggests an excellent method for testing the ion pairing formula derived in Sections VII and XI, for it

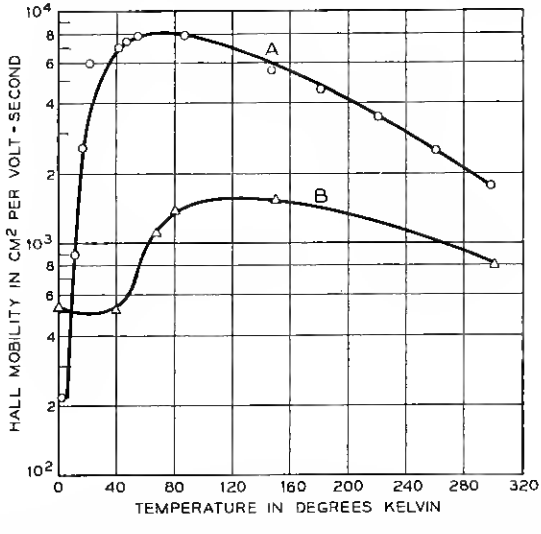

Fig. 25 — Plot of Hall mobility as a function of temperature for germanium containing 3×10^{17} cm^{-3} gallium. Curve A is for a sample containing 2.8×10^{17} cm^{-3} lithium.

enables us to determine at what temperature, at given values of N_A and N_D, P/N_D is exactly 0.5. Thus consider the fact that, all other things being equal, the control bridge and the one containing added lithium will exhibit equal Hall mobilities at a given temperature when the concentrations of charged impurities are identical in both of them. Now the concentration of such impurities in the control is simply N_A. The concentration in the bridge containing lithium is

$$N_A + N_D - 2P \tag{12.1}$$

The quantity $2P$ is removed from $N_A + N_D$, because each time a pair forms two charged scatterers are eliminated. The condition that the scattering densities in both bridges be equal is then simply

$$N_A = N_A + N_D - 2P$$

or

$$\frac{P}{N_D} = 0.5 \tag{12.2}$$

Therefore if plots of Hall mobilities versus temperature such as those appearing in Figure 25 are continued until they cross, the temperature of crossing marks the point at which P/N_D is 0.5.

In Fig. 26 typical crossings of this kind are shown. They are for two different gallium doped germanium crystals, one containing 3×10^{17} cm^{-3} gallium and the other 9×10^{15} cm^{-3}. The curves for the controls and lithium saturated samples in each case are shown as plots of the logarithm of Hall mobility against logarithm of absolute temperature. The lines plotted in this manner are straight. The lithium content of the bridge containing 9×10^{15} cm^{-3} gallium was 6.1×10^{15} cm^{-3} while that in the bridge with 3×10^{17} cm^{-3} gallium was 2.8×10^{17} cm^{-3}. All of these concentrations were obtained from Hall coefficient measurements in the controls and the lithium doped specimens.

As the temperature is increased the mobilities of the samples with lithium are reduced and approach the mobilities of the controls. This happens because pairs dissociate and more charged impurities appear. Finally when P/N_D is exactly 0.5 the curves cross. In Fig. 27 we notice that mobility measurements were not performed right up to the cross point, but that the straight lines have been extrapolated. This procedure was adopted of necessity, because of the high diffusivity of lithium. Thus, reference to Fig. 5 shows that the solubility in doped germanium decreases to a minimum as the temperature is raised from room temperature, and there is danger of precipitation. For this reason the measure-

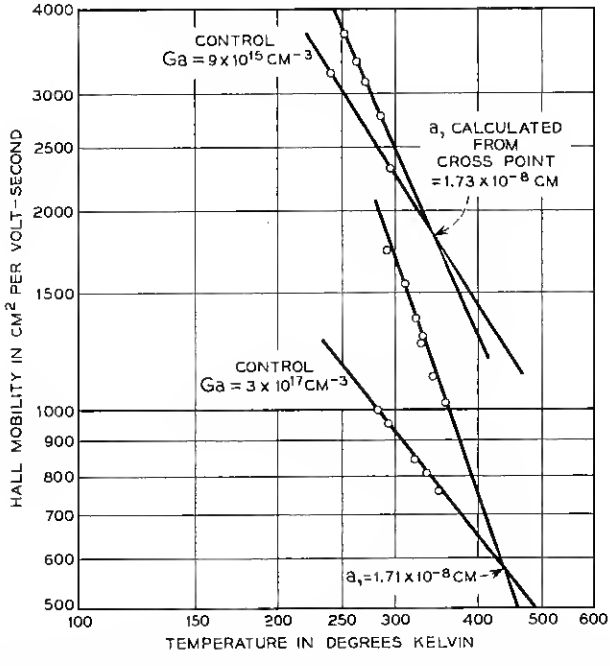

Fig. 26 — Illustration of cross over phenomenon for germanium samples containing gallium. Sample 314 contains 9×10^{15} cm^{-3} gallium and sample 302 contains 3×10^{17} cm^{-3}. Samples 316 and 301 are the corresponding samples to which lithium has been added.

ments were not carried to high temperatures.* In addition the value of the Hall coefficient was carefully checked at each temperature to see if it had changed. Since the reciprocal of the Hall coefficient[66] measures the carrier density any reduction in its value would have implied loss of compensation, or precipitation of lithium.

Over the measured points no appreciable variation of Hall coefficient was noted. Fortunately, the pairing relaxation time is quite small (less than a second) at the high temperatures involved so that it wasn't necessary to hold the samples at these temperatures for long periods in order to achieve pairing equilibrium. The times involved were too short for the occurrence of phase equilibrium characterized by precipitation.

The above discussion points up some of the care that must be taken to obtain reliable measurements. Another factor which enters the picture is the possible existence of a precipitate in the lithium doped bridge.

* In boron-doped germanium the cross-over was actually observed — no extrapolation having been necessary, because the temperature of intersection was sufficiently low.

During the course of our experiments it was discovered that precipitates have a profound effect on carrier mobility, reducing it so severely, that the mobility of the lithium doped bridge may never even rise above that of the control. Great care must be exercised in the preparation of suitable bridges to avoid the presence of precipitated lithium. Thus it may be necessary to saturate the bridge at a very low temperature (see Section IV, Figure 5) so that it is somewhat undersaturated at room temperature. This means that diffusion periods of weeks may be involved.

In Fig. 26 the sample with $N_A = 9 \times 10^{15}$ cm^{-3}, and $N_D = 6.1 \times 10^{15}$ cm^{-3} has $P/N_D = 0.5$ at 348°K, while the sample with $N_A = 3 \times 10^{17}$ cm^{-3} and $N_D = 2.8 \times 10^{17}$ cm^{-3} is half-paired at 440°K. This is to be expected, the more heavily doped specimen remaining paired up to higher temperatures. Using (9.6) and (9.3) it is possible to calculate a, the distance of closest approach of a gallium and lithium ion, from each of the measured cross points.

Thus in (9.6) we set $\theta = 0.5$, and N_A, N_D and T to correspond to each of the cases described. Having $\log_{10} Q(\alpha)$, α can be determined by interpolation in Table III and a then determined from (9.3). Of course κ is taken to be 16. Carrying through this procedure in connection with Fig. 26 leads to the satisfying result that $a = 1.71 \times 10^{-8}$ cm for the heavily doped sample and 1.73×10^{-8} cm for the lightly doped one. The values of Ω appearing in Table IV based on $a = 1.7 \times 10^{-8}$ cm therefore correspond to gallium.

Not only is this result satisfying because the two a's agree so well even though the samples involved were so different in constitution, but also because it is expected on the basis of the addition of known particle radii. Thus according to Pauling[36] the tetrahedral covalent radius of gallium is 1.26×10^{-8} cm while the ionic radius of lithium is 0.6×10^{-8} cm. Since gallium is presumably substitutional in a tetrahedral lattice we use its tetrahedral covalent radius, and since lithium is probably interstitial we use the ionic radius. The sum of the two is 1.86×10^{-8} cm which compares very favorably with the values of a quoted above.

This result constitutes good evidence that lithium is interstitial, for if it were somehow substitutional we might expect a to be something like a germanium-germanium bond length which is 2.46×10^{-8} cm. Such a value of a would lead to profoundly different crossing temperatures (of the order of 100° lower) so that it is not very likely.

One further point needs mention. This is the fact that as the two ions approach very closely, the concept of the uniform macroscopic dielectric constant, κ, loses its meaning. In fact, the binding energy should be increased (as though κ were reduced). Crude estimates of the magnitude

of this effect based on a dielectric cavity model show it to be of the order of some 10 or 15 percent of the energy computed on the assumption of the dielectric continuum, the increased binding energy showing up as a reduced value of a. This may account for the fact that the observed a, at 1.7×10^{-8} cm is less than the theoretical value, 1.86×10^{-8} cm.

The above example shows the ion pairing phenomenon in action as a structural tool, useful in investigating isolated impurities. In particular the demonstration that lithium is interstitial is interesting. The values of a have much more meaning as independent parameters in solids than they have in liquids, where a given ion may be surrounded by a sheath of solvating solvent molecules. Under the latter conditions the value of a can only be determined through application of the ion pairing theory itself.

Of course, certain unusual situations arise in solids also, and values of a (determined from ion pairing) are valuable indications of structural peculiarities.

Similar experiments have been performed on specimens doped with indium and boron. The results of all our investigations on the cross-over phenomenon are tabulated in Table V. In the table the first column lists the acceptor involved, and the second and third the appropriate concentrations of impurities. The fourth column contains the cross-over temperature, while the fifth, the measured value of a determined from it. The last column lists the values of a to be expected on the basis of the addition of tetrahedral covalent radii to the ionic radius of lithium — all of which appear in Pauling.[36]

The reliability of the measurements are in the order gallium, aluminum, boron, and indium. The principal reason for this is that the indium crystal was not grown specially for this work and was somewhat nonuniform. Of the two values obtained for a we tend to place more confi-

TABLE V

Acceptor	Acceptor conc. (cm^{-3})	Lithium conc. (cm^{-3})	Cross-over Temp. (°C.)	Measured a (cm)	Pauling a (cm)
B	7.0×10^{16}	5.9×10^{16}	338	2.05×10^{-8}	1.48×10^{-8}
B	7.0×10^{16}	5.54×10^{16}	320	2.27×10^{-8}	1.48×10^{-8}
B	7.0×10^{16}	5.85×10^{16}	330	2.16×10^{-8}	1.48×10^{-8}
Al	9.5×10^{15}	9.0×10^{15}	350	1.68×10^{-8}	1.86×10^{-8}
Ga	3.0×10^{17}	2.8×10^{17}	440	1.71×10^{-8}	1.86×10^{-8}
Ga	9.0×10^{15}	6.1×10^{15}	348	1.73×10^{-8}	1.86×10^{-8}
In	3.3×10^{17}	1.9×10^{17}	476	1.61×10^{-8}	2.04×10^{-8}
In	3.3×10^{17}	2.68×10^{17}	426	1.83×10^{-8}	2.04×10^{-8}

dence in 1.83×10^{-8} than in 1.61×10^{-8} cm. More work is necessary, however, before a real decision can be made.

A feature of Table V is the fact that gallium, aluminum, and indium exhibit orthodox behavior, i.e., the measured a's are in both cases slightly less than those expected on the basis of the addition of radii. The internal consistency of the theory gains support from the fact that gallium and aluminum behave similarly as the Pauling a's tabulated in Table V predict. In fact if 1.83×10^{-8} cm is taken as the more reliable indium value the three cases fail to match the Pauling radii by about the same amount, a result which implies that the disparity is due to the same cause, i.e., failure of the dielectric continuum concept.

Another feature of Table V is the fact that boron is out of line to the extent that the measured *a exceeds* the Pauling a by 50 per cent. A possible explanation is the following. The tetrahedral radii of boron and germanium are poorly matched (0.88 Å and 1.26 Å, respectively). The strain in the boron-germanium bond may appear as a distortion of the germanium atom in such a way as to increase the effective size of the boron ion. This strain was mentioned before in Section V where it was invoked to explain the stability of LiB^- complex in silicon.

XIII. RELAXATION STUDIES

The relaxation time discussed in Section X has been studied experimentally. The following procedure was used. A specimen was warmed to 350°K where a considerable amount of pair dissociation occurred, and then cooled quickly by plunging into liquid nitrogen. It was then rapidly transferred to a constant temperature bath, held at a temperature where pair formation took place at a reasonable rate, and the change in sample conductivity (as pairing took place) was measured as a function of time.

The principle upon which this measurement is based is the following. At a given temperature the occurrence of pairing does not change the carrier concentration, only the carrier mobility. As a result the measurement of conductivity is effectively a measurement of relative mobility. During relaxation the densities of charged impurities are changed, at the most, by amounts of the order of 50 per cent. Over this range, the mobility may be considered a linear function of scatterer density. The dependence of conductivity on time, as pairing takes place, must be of the form

$$\sigma = \sigma_\infty - \Phi\, e^{-t/\tau} \tag{13.1}$$

where σ_∞ is the conductivity when $t = \infty$, and τ is the relaxation time defined in section X while Φ is some unknown constant, depending among

other things on the initial state of the system. Equation (13.1) is based on the assumption that the number of charged scatterers decays as a first order process, and that σ is a linear function of this number, relative to the exponential dependence on time.

The first order character of pairing is fortunate for it renders the measurement of τ independent of a knowledge of Φ, i.e. independent of the initial state of the system. This is not only fortunate from the point of view of calculation but from experiment, since it is almost impossible to prepare a specimen in a well defined initial state.

The unimportance of Φ is best seen by plotting the logarithm of $\sigma_\infty - \sigma$ against time. According to (13.1) this plot is specified by

$$\log (\sigma_\infty - \sigma) = \log \Phi + \frac{t}{\tau} \tag{13.1}$$

Thus tbe reciprocal of its slope measures τ, and Φ is not involved. Fig. 27 illustrates the data for a typical run plotted in this manner. The sample is one containing about 9×10^{15} cm^{-3} gallium and the experiment was performed at 195°K (dry ice temperature). Notice that the curve is absolutely straight out to 3500 minutes, demonstrating beyond a doubt that the process is first order. The relaxatiou time computed from its slope is 1.51×10^5 seconds as against a value calculated by the methods

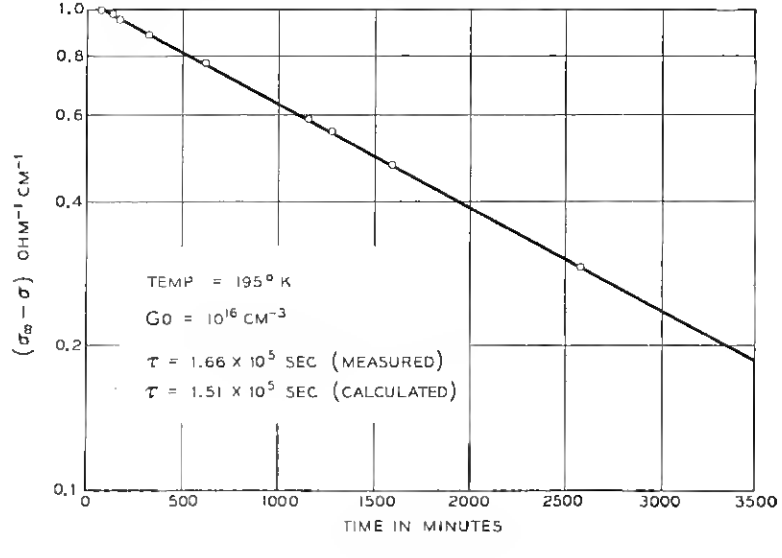

Fig. 27 — Plot of log $(\sigma_\infty - \sigma)$ as a function of time showing first order kinetics of pairing.

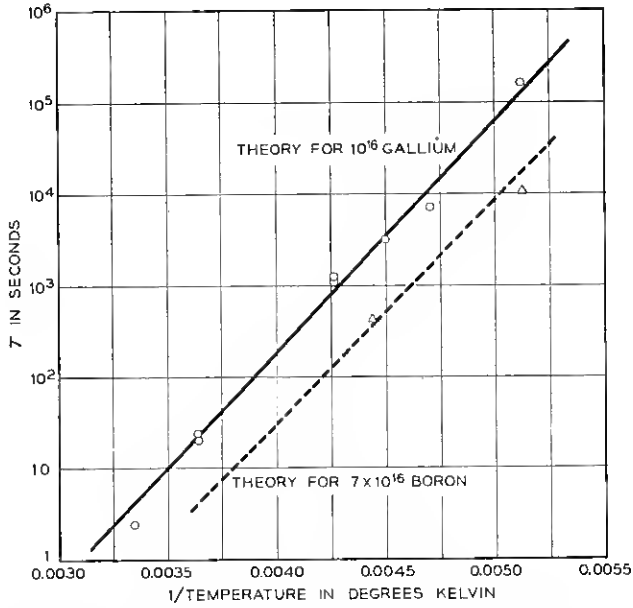

Fig. 28 — Plots of logarithm of relaxation time versus reciprocal temperature showing agreement between theory and experiment.

of section X of 1.66×10^5 seconds. The result is in good agreement with theory.

Studies of the kind illustrated in Fig. 27 have been carried out in samples doped to various levels and also at various temperatures. Boron and indium have been used as doping agents, as well as gallium. Relaxation times have been measured over the range extending from about a second to hundreds of thousands of seconds. In each case straight line plots were obtained and the agreement between calculated and measured τ's has been as good as in the example illustrated by Fig. 28. Relaxation connected with dissociation has also been measured with equally satisfactory results.

Some of these data are shown in Fig. 29 where log τ is plotted as a function of reciprocal temperature for gallium and boron at two different values of doping. The drawn curves are theoretical obtained from Fig. 20 while the points shown are experimental. It is seen that agreement is nearly perfect. The relaxation time, true to the demands of theory, does not seem to depend on the kind of acceptor used for doping, i.e., it is independent of a, the distance of closest approach.

The data in Fig. 28 actually can be used to measure the diffusivity of

lithium. As must be the case from the above mentioned agreement, the values of D_0 computed from them agree with the diffusion data of Fuller and Severiens[62] almost perfectly. This is a very quick and sensitive method (also probably exceedingly accurate) for determining diffusivities. For example the work already completed, in effect, represents the determination of diffusivities of the order of 10^{-16} cm^2/sec within a matter of an hour, and, no doubt, smaller diffusivities could be determined by doping more heavily with acceptor.

XIV. THE EFFECT OF ION PAIRING ON ENERGY LEVELS

It was predicted in Section VIII that ion pairing would drive the electronic energy states of donors and acceptors from the forbidden energy region. In this section it will be demonstrated by low temperature Hall effect measurements that the addition of lithium to gallium-doped germanium does indeed result in the removal of states from the forbidden gap rather than in the simple compensation which occurs when a non-mobile donor such as antimony is added.

At low temperatures where carrier concentration, p, is less than the donor concentration, it can be expressed in the form[67]

$$p = \frac{N_A - N_D}{N_D} \left(\frac{2\pi m_p kT}{h^2}\right)^{3/2} \exp\left[-E_A/kT\right] \tag{14.1}$$

where N_A and N_D are the concentrations of acceptor and donor states, respectively, m_p, the effective mass of free holes, h, Plank's constant, and E_A the ionization energy of the acceptor. The values of m_p and E_A are known for the group III acceptors.[68]

Lithium was added to a specimen of germanium known to contain 1.0×10^{16} cm^{-3} gallium atoms and a negligible amount of ordinary donors. Carrier concentrations for this specimen were determined from Hall coefficient measurements. The logarithm of this concentration is shown in Fig. 29 plotted against reciprocal temperature. The high temperature limit of this plot fixes $N_A - N_D$ at 1.15×10^{15} cm^{-3}.

At low temperatures the curve exhibits an extended linear portion to which (14.1) should apply. Evaluating (14.1) with $p = 4.0 \times 10^{13}$ cm^{-3} at $1/T = 0.06$ deg^{-1} and $N_A - N_D = 1.15 \times 10^{15}$ cm^{-8} we find that $N_D = 2.6 \times 10^{14}$ cm^{-3} and $N_A = 1.4 \times 10^{15}$ cm^{-3}.

Therefore, the density of apparent acceptor states has been decreased by $1.0 \times 10^{16} - 1.4 \times 10^{15} = 8.6 \times 10^{15}$ cm^{-3}. The added concentration of lithium was 1.0×10^{16} cm^{-3} $- 1.15 \times 10^{15}$ cm^{-3} $= 8.85 \times 10^{15}$ cm^{-3}, *almost identical with the loss in concentration of acceptor states.* This implies (as would be expected) that the lithium is almost totally paired.

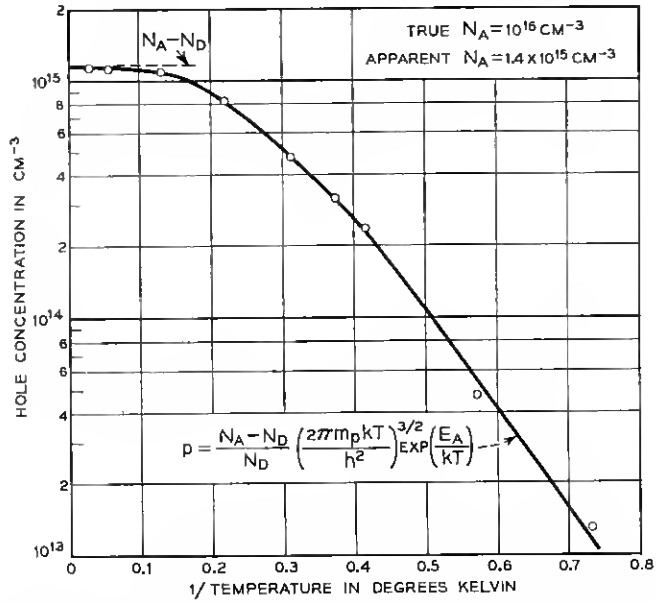

Fig. 29 — Plot of hole concentration as a function of reciprocal temperature for a sample containing ion pairs.

An even more striking result appears. From the above results the density of lithium atoms involved in pairs is 8.85×10^{15} cm^{-3} $- 2.6 \times 10^{14}$ cm^{-3} $= 8.6 \times 10^{15}$ cm^{-3}, *the same number by which the density of acceptors has been decreased!* There can be little question that ion pairing is the mechanism responsible for the removal of states.

In closing it is worth pointing out that the density of unpaired lithiums 2.6×10^{14} cm^{-3}, is certainly not characteristic of the low temperatures at which the above Hall measurements were performed. Obviously a density characteristic of some higher temperature has been quenched into the specimen. At the low temperature involved the unpaired density would be effectively zero.

XV. RESEARCH POSSIBILITIES

The fields described in the preceding text have been hardly touched, even by this long paper, and it does not seem fitting to close without some speculation concerning the possibilities of future work.

In the first place, there are other donors and acceptors besides lithium which are reasonably mobile in germanium or silicon, e.g. copper, iron,

zinc or gold. To some extent the methods of this paper can be applied to these. Furthermore, returning to lithium, there are impurities both mobile and immobile which introduce more than one energy state into the forbidden gap. The phase relations of lithium in the presence of these should be extremely interesting since the corresponding mass action equations are more complicated. Analogues of dibasic[69] acids and bases should exist.

In the case of ion pairing doubly charged acceptors like zinc in germanium[64] should be extremely interesting, since large amounts of pairing should persist up to very high temperatures. In fact such studies represent excellent means of testing for the existence of doubly charged ions. There is also the question of what happens to the two energy levels when an acceptor like zinc pairs with a single lithium ion. Are both levels driven from the forbidden gap or do they split under the perturbation?

Then there is the problem of ion *triplets* — a possibility with impurities like zinc — which is unexplored both theoretically and experimentally. Also, very strange diffusion effects must occur in the presence of doubly charged ions, to say nothing of the effect which uncompensated mobile holes might have on relaxation processes.

The field of ion pairing in silicon is relatively unexplored.

All of the phenomena discussed in this paper must occur in the group III–V compounds, more or less complicated by additional effects.

The question of the formation of the LiB^- complex in both germanium and silicon needs further study. It should behave as an acceptor and its electronic energy state might be revealed by suitable quenching techniques.

Non ionic reactions between group V donors and group III acceptors very likely occur, i.e., a real III–V covalent bond may be formed between such atoms dissolved in germanium or silicon at high temperatures. This possibility could be investigated by looking for changes in carrier mobility or impurity energy levels upon extended heating — in much the same way that ion pairing has been studied. If found, the phenomenon may provide an excellent technique for measuring the diffusitivities of all classes of impurities even at fairly low temperatures.

Such compounds may possess strange energy levels and be responsible for unexplained traps and recombination centers.

The effect of stress on the extent of ion pairing may well be profound since there will be a tendency for such stress to concentrate at imperfections. Stress studies on ion pairing may therefore be useful for further investigating the strain about an isolated impurity.

Ion pairing between lithium ions and acceptor centers located in dislocations or vacancies should occur. In the first case the dislocation

would be the analogue of the polyelectrolyte molecule in the aqueous solution.

An interesting question, in the diffusion of substitutional acceptors, concerns whether the ion or the neutral atom is responsible for diffusion. It is possible that the neutral atom, less securely bonded to the lattice, is the chief agent. This might be determined by changing the ratio of neutral atoms to ions by suitably doping with other donors or acceptors.

Doping apparently affects the concentration of vacancies which have acceptor properties and therefore the rate of diffusion.[70, 71]

Other interesting effects concerning the distribution of an impurity between two different kinds sites in the lattice[72] are also possible.

These and many other fascinating fields still require exploration. We hope to investigate some of them in the near future.

ACKNOWLEDGMENTS

The authors are greatly indebted to A. J. Pietruszkiewicz, Jr., for assistance in carrying out experimental work relating to solubility and diffusion and to J. P. Maita for help with experimental work on Hall effect and an ion-pair relaxation. Thanks are due N. B. Hannay for many helpful comments during the course of the work and during preparation of the manuscript. Thanks are also due Miss M. C. Gray for the evaluation of the integrals in Section VII and to F. G. Foster for the photograph of Fig. 8. Finally the authors would like to thank the editors of the Bell System Technical Journal for providing space so that all of the important features of our subject could be treated in one article.

Appendix A

THE EFFECT OF ION PAIRING ON SOLUBILITY

In Section VIII attention was called to the fact that ion pairing should have some effect on lithium solubility but that it would be difficult to achieve conditions under which the effect would be observable. Now, this point will be enlarged upon. Consider an equilibrium like (2.1) except imagine it to take place in germanium with gallium as the immobile acceptor. (This because germanium with gallium has been studied in ion pairing investigations.)

$$\begin{array}{ccccc} \text{Li (external)} \rightleftarrows & \text{Li}^+ & + & e^- & \\ & + & & + & \\ & \text{Ga}^- & + & e^+ & \\ & \Updownarrow & & \Updownarrow & \\ & [\text{Li}^+\text{Ga}^-] & & e^+e^- & \end{array} \qquad \text{(A1)}$$

where $[Li^+Ga^-]$ represents an ion pair, whose concentration we denote by P. N_A and N_D will be the total densities of acceptor and donor respectively and A^- and D^+ the densities of acceptor and donor ions in the unpaired state.

As in the main text, n and p will represent the concentrations of holes and electrons. The following relations are then to be expected on the basis of definition, mass action, and charge balance.

$$N_A = A^- + P \tag{A2}$$

$$N_D = D^+ + P \tag{A3}$$

$$D^+ n = K^* \tag{A4}$$

$$np = n_i^2 \tag{A5}$$

$$\frac{P}{A^+ D^-} = \Omega \tag{A6}$$

$$D^+ + p = A^- + n \tag{A7}$$

Equations (A4), (A5), and (A7) are just reproductions of (3.1), (3.2), (2.8), while (A6) is the same as (9.4). The problem is to express the solubility of lithium, N_D, as a function of N_A. Manipulation of the preceding set of equations gives this result as

$$N_D = \frac{(N_A - A^-)(1 + \Omega A^-)}{\Omega A^-} \tag{A8}$$

with A^- given by the solution of

$$\frac{N_A - A^-}{\Omega A^-} = \frac{A^-}{1 + \sqrt{1 + \left(\frac{2n_i}{D_{0+}}\right)^2}} + \sqrt{\left(\frac{A^-}{1 + \sqrt{1 + \left(\frac{2n_i}{D_0^+}\right)^2}}\right)^2 + (D_0^+)^2} \tag{A9}$$

where D_0^+ is defined by (3.3). Equation (A9) generally needs to be solved numerically for A^-.

To see what these relations predict in a special case consider the solubility of lithium in gallium-doped germanium at 300°K. At this temperature the values of n_i and D_0^+ and Ω are

$$\begin{aligned} n_i &= 2.8 \times 10^{13}\ \text{cm}^{-3} \\ D_0^+ &= 7 \times 10^{13}\ \text{cm}^{-3} \\ \Omega &= 1.61 \times 10^{15}\ \text{cm}^{-3}. \end{aligned} \tag{A10}$$

TABLE AI — TEMPERATURE = 300°K

N_A (cm^{-3})	N_D (cm^{-3})	N_D^* (cm^{-3})	$P = N_A - A^-$ (cm^{-3})
10^{14}	1.25×10^{14}	1.25×10^{14}	0.15×10^{14}
10^{15}	0.94×10^{15}	0.875×10^{15}	0.44×10^{15}
10^{16}	0.985×10^{16}	0.875×10^{16}	0.77×10^{16}
10^{17}	0.990×10^{17}	0.875×10^{17}	0.92×10^{17}
10^{18}	0.995×10^{18}	0.875×10^{18}	0.97×10^{18}

The value of n_i is taken from Figure 2, of D_0^+, from Figure 5, and of Ω, from Table IV. Using (A10) together with (A9) and (A8) leads to the results tabulated in Tahle AI. In this tahle, N_D^* represents the solubility for the case $\Omega = 0$, i.e., the soluhility if there were no ion pairing. The main feature to be obtained from the Tahle is that N_D is not very much larger than N_D^*, no matter how large the value of N_A. This is true in spite of tbe fact that the last column which lists P shows that at $N_A = 10^{18}$ cm^{-3} P is about 98% of N_D so that pairing of the donor is virtually *complete*.

The result is not limited to the special conditions of doping and temperature chosen in compiling Table AI, but must be quite general. One can arrive at this conclusion in the following way.

By subtracting (A3) from (A2) we obtain

$$N_A - N_D = A^- - D^+. \tag{A11}$$

The quantities A^- and D^+ appear in equations (A4) and (A7), while n and p, appearing in (A4) and (A7) are related hy (A5). These three equations are sufficient for the determination of D^+ in its dependence on A^-. That this is the case is immediately ohvious when (A4), (A5), and (A7) are recognized as reproductions of (3.1), (3.2) and (2.8). In fact this means that the desired relationship hetween D^+ and A^- is nothing more than equation (3.4) which itself is predicated on (3.1), (3.2), and (2.8). But then it is known according to (3.6), that D^+ can at the most be slightly greater than A^-, although most likely less. This assumes of course that we deal with dopings sufficiently high so that (3.5) applies. On the other hand at low dopings (3.4) tells us that D^+ will he D_0^+. Therefore if we work with a system in which in the absence of pairing the electron-hole equilibrium has driven the value of N_D close to N_A (as it has in this system — see N_D^*) the introduction of pairing cannot drive it much higher, since according to (A11) if D^+ cannot get higher than A^-, N_D cannot exceed N_A. This is evident in Tahle AI where N_D comes very close to N_A but never exceeds it.

When N_A is very small so that D^+ equals D_0^+ and does exceed A^- hy

a large amount, there can he no visible increment in solubility as a result of pairing because P can never exceed N_A which by definition is small.

The physical reason for these limitations is the following. Suppose N_D is driven close to N_A by the hole-electron equilihrium so that in terms of carriers (holes and electrons) the specimen is very closely compensated. Then if by the formation of pairs, additional donors are caused to enter the crystal, the electrons they donate cannot be ahsorbed by holes because very few of the latter are present. Thus the following two sketched equilibria will oppose each other

$$\begin{array}{ccccc} \text{Li (external)} & \rightleftarrows & \text{Li}^+ & + & e^- \\ & & + & & \\ & & \text{Ga}^- & & \\ & & \Updownarrow & & \\ & & [\text{Li}^+\text{Ga}^-] & & \end{array} \qquad \text{(A12)}$$

the one involving electrons attempting to drive lithium out of solution because of the huild-up of electron concentration, and the pairing equilibrium attempting to hring lithium into solution in order to form pairs. Thus the pairing process will not he as efficient a solubilizer as might be thought at first.

This point can he illustrated hy considering a situation in which the germanium crystal not only contains gallium to the level, N_A but also an immobile donor, to the level $N = 0.99\,N_A$. Thus, tbe crystal is almost compensated before any lithium has been added. Nevertheless, there are still N_A gallium ions so that even though the hole-electron equilibrium, working on the differential, $0.01\,N_A$, cannot increase the solubility of lithium, the pairing process might. To investigate this situation equations (A2) to (A7) can he adopted with the simple change that $(A^- - N)$ replaces A^- in (A7).

Taking the situation covered by (A10) at 300°K, Tahle AII was eompiled. Here again $N_D{}^*$ is the solubility for $\Omega = 0$.

If only the hole-electron effect were operative, then we could not expect to drive N_D much heyond $N_A - N$. In the 10^{16} case $N_A - N$ is 10^{14} cm^{-3} and in the 10^{17} case it is 10^{15} cm^{-3}. The values of $N_D{}^*$ in Table AII thus eonfirm this argument. Furthermore, N_D is in neither case much greater than $N_D{}^*$ showing that despite the fact that there were, respec-

TABLE AII — TEMPERATURE 300°K

N_A (cm^{-3})	N (cm^{-3})	N_D (cm^{-3})	$N_D{}^*$ (cm^{-3})	P (cm^{-3})
10^{16}	0.99×10^{16}	3.2×10^{14}	1.26×10^{14}	3×10^{14}
10^{17}	0.99×10^{17}	1.6×10^{16}	0.88×10^{15}	1.6×10^{16}

tively, 10^{16} and 10^{17} cm^{-3} gallium ions available for pairing, the pairing process did very little to increase the solubility.

If the constant Ω is exceedingly large as is probably the case for a multiply charged acceptor, it is possible that ion paring will have a measurable effect on solubility.

Appendix B

Concentration dependence of diffusivity in the presence of ion pairing

In Section VIII it was mentioned that the diffusivity of a mobile donor like lithium is concentration dependent when the donor participates in a pairing equilibrium with an immobile acceptor. In this appendix we propose to investigate the nature of the dependence.

Consider a semiconductor, uniformly doped to the level, N_A, with acceptor. Let the local density of mobile donor be $N_D(x)$, x being the position coordinate. If $P(x)$ is the local pair concentration, then the local density of free diffusible ions is $(N_D - P)$. The flux of these diffusing ions then depends upon the gradient (assuming Fick's law[49]) of $(N_D - P)$. Thus, if D_0 is the diffusivity of free donor, i.e. the diffusivity in the absence of pairing, then the flux density is

$$f = -D_0 \frac{\partial (N_D - P)}{\partial x} \tag{B1}$$

If we apply (9.4) to the present case we can write

$$\Omega = \frac{P}{(N_A - P)(N_D - P)} = \frac{-(N_D - P) + N_D}{[(N_A - N_D) + (N_D - P)](N_D - P)} \tag{B2}$$

from which it is possible to solve for $(N_D - P)$. Thus

$$N_D - P = \frac{1}{2}\left(N_D - N_A - \frac{1}{\Omega}\right) + \sqrt{\frac{1}{4}\left(N_D - N_A - \frac{1}{\Omega}\right)^2 + \frac{N_D}{\Omega}} \tag{B3}$$

Substitution of (B3) into (B1) yields

$$f = -\frac{D_0}{2}\left[1 + \frac{\frac{1}{2}\left(N_D - N_A + \frac{1}{\Omega}\right)}{\sqrt{\frac{1}{4}\left(N_D - N_A - \frac{1}{\Omega}\right)^2 + \frac{N_D}{\Omega}}}\right]\frac{\partial N_D}{\partial x} \tag{B4}$$

If ion pairing was not thought of, the flux density would have been written in terms of the gradient of the total concentration, N_D.

$$f = -D\frac{\partial N_D}{\partial x} \tag{B5}$$

where D is the diffusivity. Comparison of (B5) with (B4) leads to the relation

$$D = \frac{D_0}{2}\left[1 + \frac{\frac{1}{2}\left(N_D - N_A + \frac{1}{\Omega}\right)}{\sqrt{\frac{1}{4}\left(N_D - N_A - \frac{1}{\Omega}\right)^2 + \frac{N_D}{\Omega}}}\right] \tag{B6}$$

so that D depends on the local concentration, N_D, of diffusant.

It is interesting to explore the limiting forms of D when $N_D \ll N_A$ and when $N_D = N_A$. In the latter case (B6) reduces to

$$D = \frac{D_0}{2}\left[1 + \frac{\frac{1}{2\Omega}}{\sqrt{\frac{1}{4\Omega^2} + \frac{N_A}{\Omega}}}\right] \tag{B7}$$

wbile (B3) becomes

$$N_A - P + \frac{1}{2\Omega} = \sqrt{\frac{1}{4\Omega^2} + \frac{N_A}{\Omega}}\,. \tag{B8}$$

Substituting the left side of (B8) for the denominator involving the radical in (B7) leads to

$$D = \frac{D_0}{2}\left[1 + \frac{1}{2(N_A - P)\Omega + 1}\right] \tag{B9}$$

But according to (B2), when $N_A = N_D$,

$$(N_A - P)\Omega = \frac{P}{N_A - P} \tag{B10}$$

so that (B9) becomes

$$D = \frac{D_0}{2}\left[1 + \frac{1}{\frac{2P}{N_A - P} + 1}\right] \tag{B12}$$

Now in case the degree of pairing is higb (wbicb is, of course, the case we are interested in) P will be almost equal to N_A so tbat

$$\frac{2P}{N_A - P} \tag{B13}$$

will be a very large number. If this is so the second term in brackets in (B12) can be set equal to zero and we have

$$D = \frac{D_0}{2}\,. \tag{B14}$$

In the other extreme with $N_D \ll N_A$ (B6) becomes

$$D = \frac{D_0}{2}\left[1 + \frac{\frac{1}{2}\left(\frac{1}{\Omega} - N_A\right)}{\sqrt{\frac{1}{4}\left(\frac{1}{\Omega} + N_A\right)^2}}\right] = \frac{D_0}{1 + \Omega N_A} \tag{B15}$$

Since ΩN_A can exceed unity by a large amount it is evident that the relation in (B15) predicts a large reduction in diffusivity towards the front end of a diffusion curve where $N_D \ll N_A$, and (B14) a smaller reduction in D_0 where N_D may be close to N_A. That part of the medium near the front of the diffusion curve acts therefore like a region of high resistance, confining the diffusant to the back end where the resistance is low.

Appendix C

Solution of boundary value problem for relaxation

In Section X equations (10.23), (10.21), (10.20), and (10.19) defined a boundary value problem which we reproduce here, except that (10.20) and (10.19) have been written more completely with the aid of (10.16). Thus

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial \rho}{\partial r} + R\rho\right) = \frac{1}{D_0}\frac{\partial \rho}{\partial t} \tag{C1}$$

$$\frac{\partial \rho}{\partial r} + \frac{R}{r^2}\rho = 0, \qquad r = L, \qquad r = a \tag{C2}$$

$$\rho = N^2, \qquad t = 0, \qquad a < r < L \tag{C3}$$

In principle this problem is soluble by separation of variables.[66] Thus we define

$$\rho(r, t) = G(r)\, S(t) \tag{C4}$$

which upon substitution into (C1), yields the two ordinary differential equations

$$\frac{d}{dr}\left[r^2\frac{dG}{dr} + RG\right] + \eta^2 G = 0 \tag{C5}$$

$$\frac{d\ \ell n\ S}{dt} + \eta^2 D_0 = 0 \tag{C6}$$

where η^2 is an arbitrary positive parameter.

The allowable values of η are determined by (C2) which can now be

replaced by

$$\frac{dG}{dr} + \frac{R}{r^2} G = 0, \qquad r = L, \qquad r = a \tag{C7}$$

Equation (C6) can be solved immediately to give

$$S_\eta(t) = e^{-\eta^2 D_0 t} \tag{C8}$$

and if we assign the subscript η to the G going with η the most general solution of (C1) and (C2) will be

$$\rho = \sum_\eta A_\eta G_\eta(r) e^{-\eta^2 D_0 t} \tag{C9}$$

where the A_η are arbitrary constants so determined that (C3) is satisfied.

Equation (C9) shows that in reality there exists, for this problem, a spectrum of relaxation times, $1/\eta^2 D_0$. After a brief transient period many of the higher order terms will decay away and eventually only the first two terms will have to be considered. Finally when equilibrium is attained only the first term will survive.

The last statement implies that $\eta = 0$, is an allowable eigenvalue, i.e., that the first term is independent of time. That this is so can be proved by solving (C5) for $\eta = 0$, and substituting the result in (C7). Thus

$$G_0(r) = \exp\left(\frac{R}{r}\right) \tag{C10}$$

and this does satisfy (C7). ρ can then be approximated after the transient by

$$\rho = A_0 \exp\left(\frac{R}{r}\right) + A_1 G_1(r) e^{-\eta_1^2 D_0 t} \tag{C11}$$

from which it is obvious that the relaxation time dealt with in section X is

$$\tau = \frac{1}{\eta_1^2 D_0} \tag{C12}$$

In principle it should be possible to evaluate G_1 by the straightforward solution of (C5) and determination of the second eigenvalue through substitution of this solution in (C7). In fact this represents a rather unpleasant task since G is a confluent hypergeometric function.[57] Therefore we shall follow an alternative route based on the assumption that by the time (C11) applies the flux $4\pi r^2 J^*(r)$, where J^* is given by (10.16), is almost independent of r. The reader is referred to some related papers[58, 59]

for the justification of this view. Briefly it is permissible, after a short transient period, in spherical diffusion, whenever the dimensions of the diffusion field are large compared to the dimension of the sink. This results from the fact that in spherical diffusion from an infinite field[48] a real steady state is reached after a brief transient period. In contrast, in plane-parallel diffusion to a sink from an infinite field,[60] a steady state is never reached.

Substituting (C11) into (10.16) then yields

$$J^* = -D_0 A_1 e^{-\eta_1^2 D_0 t} \left(\frac{dG_1}{dr} + \frac{R}{r^2} G_1 \right) \tag{C13}$$

Multiplying J^* by $4\pi r^2$ and demanding that the product be independent of r, leads to the relation

$$r^2 \frac{dG_1}{dr} + RG_1 = \delta \tag{C14}$$

where δ is constant. The solution of (C14) is

$$G_1 = \exp\left(\frac{R}{r}\right) + \frac{\delta}{R} \tag{C15}$$

This is a sufficient approximation for G_1.

The constants η_1, A_0, A_1, and δ must now be determined. To accomplish this we note that (C2) which specifies that the boundaries at $r = a$ and $r = L$, are impermeable is equivalent to the condition that ions be conserved with the interval (a, L), or that

$$4\pi \int_a^L r^2 \rho \, dr = N \tag{C16}$$

After infinite time ρ is specified by the first term of (C11) and when this is inserted into (C16) the result is

$$A_0 = NM \tag{C17}$$

where M is defined by (10.26).

Substitution of (C17) and (C15) into (C11) gives

$$\rho = NM \exp(R/r) + \left(A_1 \exp(R/r) + \frac{A_1 \delta}{R} \right) e^{-\eta_1^2 D_0 t} \tag{C18}$$

Now (C3) applied to (C18) demands

$$NM + A_1 = 0 \tag{C19}$$

$$\frac{A_1 \delta}{R} = N^2 \tag{C20}$$

Of course this presumes that the approximation contained in (C18) is valid down to very small values of time. This assumption is well founded as the transient does vanish after a rather short time.

Inserting (C19) and (C20) in (C18) then gives us

$$\rho = NM \exp (R/r) + N[N - M \exp (R/r)]e^{-\eta_1^2 D_0 t} \tag{C21}$$

in which only η_1 remains to be determined.

Substitution of (C21) into (C16), recalling the definitions of M and L, shows that it already satisfies (C16) for any time, t. Thus (C16) cannot be used for determining η_1.

On the other hand we note from (C21) that as soon as r becomes of order, R, ρ becomes almost independent of r, being given

$$\rho = N\{N + (N - M)e^{-\eta_1^2 D_0 t}\} \tag{C22}$$

Since L is of the order 10R or greater, this means that throughout most of the volume, $1/N$ (in fact throughout 0.999 $1/N$) ρ is independent of r. Effectively, the entire volume $1/N$ has been drained of ions, i.e., they have been trapped. The total ion content at time t, may then be taken as the product of ρ, given by (C22), with $1/N$, that is,

$$N + (N - M)e^{-\eta_1^2 D_0 t} \tag{C23}$$

The time rate of change of this content must be given by the flux $4\pi r^2 J^*$.

$$\begin{aligned} \frac{d}{dt}[N + (N - M)e^{-\eta_1^2 D_0 t}] \\ = -\eta_1^2 D_0 (N - M)e^{-\eta_1^2 D_0 t} = 4\pi r^2 J^*(r, t) \\ = -4\pi R N^2 D_0 e^{-\eta_1^2 D_0 t} \end{aligned} \tag{C24}$$

in which (C21) has been substituted into (10.16) to pass from the third to the fourth expression. Comparing the second and fourth term of (C24) reveals

$$\eta_1^2 D_0 = \frac{4\pi R N^2 D_0}{(N - M)} = \frac{4\pi q^2 N^2 D_0}{\kappa k T (N - M)} \tag{C25}$$

or

$$\tau = \frac{1}{\eta_1^2 D_0} = \frac{\kappa k T (N - M)}{4\pi q^2 N^2 D_0} \tag{C26}$$

the value quoted in (10.25).

APPENDIX D

MINIMIZATION OF THE DIFFUSION POTENTIAL

In Section V the statement was made that equation (11.2) was a valid approximation everywhere within a p type region, provided that N_D did not fluctuate through ranges of order N_A in shorter distances than

$$\ell = \sqrt{\frac{\pi\kappa kT}{q^2 N_A}} \tag{D1}$$

This statement will now be proved.

The electrostatic potential is determined by the space charge equation[31]

$$\frac{d^2V}{dx^2} = -\frac{4\pi q}{\kappa}\left[N_D(x) + p(x) - N_A\right] \tag{D2}$$

where we assume that the material is everywhere p-type so that the electron density, n, does not enter the right side of (D2). Furthermore, the mobility of holes is so much greater than that of donor ions that the former may be considered to always be at equilibrium with respect to the distribution of the latter. Boltzmann's law[29] may then be applied to p. The result is

$$p = N_A \exp\left[-qV/kT\right] \tag{D3}$$

where the potential is taken to be zero when $p = N_A$.

Choose an arbitrary point, x_0, where the potential is V_0 and investigate (D2) in its neighborhood. We wish to determine the conditions under which the right side of (D2) may be approximated by zero, i.e., the "no-space-charge condition," in this neighborhood. The limits of the neighborhood will be defined such that

$$|V - V_0| = |u| \leqq kT/2q \tag{D4}$$

so that, in it, the exponential in (D3) can be linearized

$$p = N_A \exp\left[-qV_0/kT\right]\left(1 - \frac{qu}{kT}\right) \tag{D5}$$

Then (D2) becomes

$$\frac{d^2u}{dx^2} = \frac{4\pi q}{\kappa}\left\{N_A\left[1 - \exp\left(-qV_0/kT\right)\right] - N_D(x) + \left[\frac{qN_A}{kT}\exp\left(-qV_0/kT\right)\right]u\right\} \tag{D6}$$

The no space charge condition in the defined region is therefore

$$u = \left(\frac{kT \exp (qV_0/kT)}{qN_A}\right) N_D(x) + \left(\frac{kT}{q}\right) \frac{\exp (- qV_0/kT) - 1}{\exp (- qV_0/kT)} \quad \text{(D7)}$$

To simplify notation define

$$\exp[-qV_0/kT] = \gamma_0 \quad \text{(D8)}$$

Next expand both N_D and u in Fourier series

$$N_D = \sum_{s=0}^{\infty} A_s \sin sx + B_s \cos sx \quad \text{(D9)}$$

$$u = \sum_{s=0}^{\infty} \alpha_3 \sin sx + \beta_s \cos sx \quad \text{(D10)}$$

Substitution of (D9) and (D10) into (D6) and equating coefficients of like terms leads to the set of relations

$$\beta_0 = \frac{kT}{qN_A\gamma_0} [N_A(\gamma_0 - 1) + B_0] \quad \text{(D11)}$$

$$\alpha_s = \frac{4\pi q}{\kappa} \left(\frac{\ell^2/4\pi^2\gamma_0}{1 + (s^2\ell^2/4\pi^2\gamma_0)}\right) A_s \quad \text{(D12)}$$

$$\beta_s = \frac{4\pi q}{\kappa} \left(\frac{\ell^2/4\pi^2\gamma_0}{1 + (s^2\ell^2/4\pi^2\gamma_0)}\right) B_s \quad \text{(D13)}$$

Now the wavelength of the sth component in (D9) is

$$\lambda_s = 2\pi/s \quad \text{(D14)}$$

If N_D contains no important components of wavelength shorter than

$$\frac{\ell}{\sqrt{\gamma_0}} \quad \text{(D15)}$$

the B_k for such components may be set equal to zero. But then the only terms which appear in (D12) and (D13) are terms where the denominators which (with the aid of (D14)) may be written as

$$\kappa \left(1 + \frac{\ell^2}{\gamma_0\lambda_s^2}\right) \quad \text{(D16)}$$

may be set equal to κ. Thus we have in place of (D12) and (D13)

$$\alpha_s = \frac{q\ell^2}{\kappa\pi} A_s = \frac{kT}{qN_A\gamma_0} A_s \quad \text{(D17)}$$

$$\beta_s = \frac{q^2\ell^2}{\kappa\pi} B_s = \frac{kT}{qN_A\gamma_0} B_s \tag{D18}$$

The requirement that N_D contain no Fourier terms of wavelength shorter than (D15) is obviously the condition that N_D never pass from its maximum to its minimum value in a distance shorter than D(15). Since we are assuming that N_D may at places be of order N_A, and at others, of order zero, this amounts to the condition that N_D does not fluctuate over ranges comparable with N_A in distances shorter than (D15).

The use of (D11), (D17), and (D18) in (D10) yields

$$\begin{aligned} u &= \frac{kT}{qN_A\gamma_0}\left[N_A(\gamma_0 - 1) + \sum_{s=0}^{\infty}(A_s \sin sx + B_s \cos sx)\right] \\ &= \frac{kT}{q}\frac{(\gamma_0 - 1)}{(\gamma_0)} + \frac{kT}{q\gamma_0}\frac{N_D}{N_A} \end{aligned} \tag{D19}$$

which by reference to the definition (D8) for γ_0 proves to be identical with (D7), the no-space-charge condition.

Equation (D19) is only true when N_D does not fluctuate through ranges of order, N_A, in distances smaller than $\ell/\sqrt{\gamma_0}$. This distance depends on γ_0 and thus on the point where $V = V_0$, whose neighborhood is being explored. Thus, we may say that there will be no space charge at all points whose V_0 is such as to fix γ_0 at a value such that

$$\gamma_0 > \frac{\ell^2}{\lambda_{\min}^2} \tag{D20}$$

where $\lambda_{\min}$ is the minimum wavelength which needs to be considered in the Fourier expansion of N_D. In terms of the definition of γ_0 this means

$$V_0 < \frac{kT}{q}\, \ell n\, \frac{\lambda_{\min}^2}{\ell^2} \tag{D21}$$

Thus, at all points where V_0 is less than the right side of (D21) the no space charge approximation will hold. (D21) shows, that in the limit when $\lambda_{\min}$ goes toward zero, i.e. when the infinite series must be used for N_D, the right side of (D21) will approach $-\infty$ and V_0 will satisfy (D21) hardly anywhere. Thus space charge will exist almost everywhere.

In most diffusion problems the extremes of potential will occur in regions where there is no space charge. Thus in one extreme N_D may equal $0.9\, N_A$ and in the other it may equal zero. If there is no space charge in these extremes we may write for them

$$N_A - N_D = p = N_A \exp(-qV/kT) \tag{D22}$$

in which (D3) has been used. Setting N_D equal to zero in one extreme

yields $V = 0$. In the other extreme $N_D = 0.9\ N_A$ so that we get

$$V = \frac{kT}{q}\ \ell n\ 10 \tag{D23}$$

This therefore is the largest value which V_0 may assume in our case. Inserting the expression in D21 in place of V_0 we end with the relation

$$10 < \frac{\lambda^2_{\min}}{\ell^2} \tag{D24}$$

Thus provided that in the distribution being considered

$$\lambda_{\min} > 3.5\ell \tag{D25}$$

there will be no space charge anywhere.

At high temperatures $0.1\ N_A$ may be less than n_i. Under these conditions (D24) should be replaced by

$$\frac{N_A}{n_i} < \frac{\lambda^2_{\min}}{\ell^2} \tag{D26}$$

and in the limit that n_i becomes very large it is obvious that (D26) will always be satisfied. The rule to be enunciated for the cases we shall be interested in is the one given in section XI, i.e. that no space charge will exist provided that $\lambda^2_{\min}$ is no less than order, ℓ.

Appendix E

Calculation of Diffusivities from Conductances of Diffusion Layers

In this appendix equation (11.12) will be derived. In the first place we note that the dependence of N_D on position x, and time t, will be of the form $N_E(x/\sqrt{t})$ at any stage of the diffusion process. This results from a theorem due to Boltzmann[61] that when the dependence of D upon x and t is of the form $D(N_D)$, i.e., the dependence is through N_D, and a semi-infinite region extending from $x = 0$ to $x = \infty$ is being considered, then, in the case of plane parallel diffusion, the only variable in the problem will be $x/\sqrt{t}$.

Although the wafers considered in Section XI are of finite thickness d, the stages of diffusion investigated are such that the two regions of loss near the surfaces have not contacted each other. As a result the system behaves like two semi-infinite regions backed against one another, and the preceding arguments hold. The conductance Σ, defined in section XI will be proportional to the integral of the product of the local carrier

density by the local mobility. Thus

$$\Sigma = \omega \int_0^{d/2} \mu(x, t)[N_A - N_D(x, t)]\, dx \tag{E1}$$

where ω is a proportionality constant and $\mu(x, t)$ is the local mobility. An upper limit of $d/2$ rather than d is used because of symmetry. The local mobility will vary because N_D, and therefore the local density of charged impurity scatterers,[54] varies. Let $N_D{}^0$ be the initial uniform density (before any diffusion out) of donors, and write (E1) as

$$\begin{aligned}\Sigma &= \omega \int_0^{d/2} \mu(x, t)[N_A - N_D{}^0 + N_D{}^0 - N_D(x, t)]\, dx \\ &= \omega \int_0^{d/2} \mu(x, t)[N_A - N_D{}^0]\, dx + \omega \int_0^{\infty} \mu(x, t)[N_D{}^0 \\ &\qquad - N_D(x, t)]\, dx\end{aligned} \tag{E2}$$

The second integral on the right of (E2) is given the upper limit ∞, because in the experiments we wish to perform $N_D{}^0 - N_D$ becomes zero long before x reaches $d/2$.

Now in the first integral on the right of (E2) we may set $\mu(x, t)$ equal to the constant value μ_0, which it assumes in the bulk of the wafer, because the breadtb of the depletion layer near the surface (in which $\mu(x, t)$ departs from μ_0) is small compared to $d/2$. The same thing cannot be done in the second integral since the integrand vanishes beyond the depletion layer and the total contribution *comes* from that layer. We thus obtain

$$\begin{aligned}\Sigma &= \omega\mu^0(N_A - N_D{}^0)\, d/2 \\ &\quad + \omega \int_0^{\infty} \mu\left(\frac{x}{\sqrt{t}}\right)\left[N_D{}^0 - N_D\left(\frac{x}{\sqrt{t}}\right)\right] dx\end{aligned} \tag{E3}$$

In the integral in (E3) both μ and N_D are represented as functions of $x/\sqrt{t}$, the latter because of what has been said above, and the former, because it is a function of the latter. Defining

$$\nu = x/2\sqrt{Dt} \tag{E4}$$

in which D is constant, and substituting in (E3) gives finally

$$\Sigma = \omega\mu_0(N_A - N_D{}^0)d/2 + 2\omega\sqrt{Dt} \int_0^{\infty} \mu(\nu)[N_D{}^0 - N_D(\nu)]d\nu \tag{E5}$$

Since the definite integral is a constant (E5) shows that Σ is a linear function of $\sqrt{t}$, a fact mentioned in section XI.

In order to make use of the measured dependence of Σ on $\sqrt{t}$ to determine diffusivities, the functions $\mu(\nu)$ and $N_D(\nu)$ must be specified. For the latter we shall assume the Fick's law solution[62]

$$N_D = N_D^0 \operatorname{erf} \nu \tag{E6}$$

going with constant D, and $N_D = 0$ as a boundary condition at $x = 0$ at the surface. (In section XI the limitations of this assumption in the presence of ion pairing and diffusion potential are discussed.) The ν dependence of μ is more complicated. In general, we shall be concerned with electrical measurements in two extreme cases. In the first case ion pairing, under the condition of measurement, is everywhere complete so that the local density of scatterers will be given by

$$N_A - N_D(\nu) \tag{E7}$$

In the other case ion pairing will be entirely absent, so that the local scatterer density, will be specified by

$$N_A + N_D(\nu) \tag{E8}$$

In all experiments N_A will be only slightly greater than N_D^0 so that it may be replaced by this quantity. Doing this, and substituting (E6) into (E8) and (E9) gives

$$N_D^0 \operatorname{erfc} \nu = N(\nu) \tag{E9}$$

for the scattering density in the ion pairing case, and

$$N_D^0(1 + \operatorname{erf} \nu) = N(\nu) \tag{E10}$$

for the no pairing case.

Since almost all our experiments have been in germanium we now specialize our attention to that substance. However, the procedure invoked below can be applied to silicon as well.

The dependence of hole mobility, μ, on scattering density, N, for germanium at room temperature is shown in Fig. 30 taken from Prince's data.[63] The integral in (E5) assumes the form

$$\mathrm{N}_D^0 \int_0^\infty \mu(\mathrm{N}(\nu)) \operatorname{erfc} \nu \, d\nu. \tag{E11}$$

Choosing $N(\nu)$ as either (E9) or (E10) and using Fig. 30 together with a table of error functions makes the numerical evaluation of (E11) possible. Since $N(\nu)$ given by (E9) or (E10) depends on N_D^0, so will the integral.

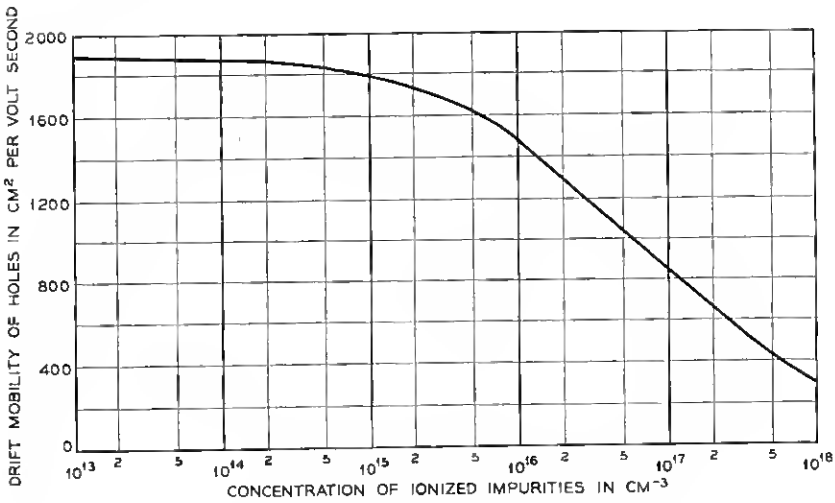

Fig. 30 — Plot of hole-drift mobility in germanium as a function of ionized impurity concentration after Prince.

The numerical evaluation has been performed for a range of N_D^0 in both the pairing and non-pairing cases. In this manner it has been possible to evaluate the "correction factor" ϑ defined by the following equation

$$\int_0^\infty \mu(\nu)\ \text{erfc}\ \nu\ d\nu = \vartheta\mu_\infty \int_0^\infty \text{erfc}\ \nu\ d\nu \qquad \text{(E12)}$$

$$= \vartheta\mu_\infty(0.563)$$

where μ_∞ is the mobility in the presence of N_A scatterers. Fig. 22 contains plots (for germanium) of $\vartheta(N_D^0)$ versus N_D^0 for both the pairing and non-pairing cases. It is seen that ϑ is never much different from unity.

Equation (E5) can now be written as

$$\Sigma = \omega\mu_0(N_A - N_D^0)\,d/2 + \omega\mu_\infty[1.128\vartheta N_D^0\sqrt{D}]\sqrt{t} \qquad \text{(E13)}$$

Defining

$$\Sigma_0 = \omega\mu_0(N_A - N_D^0)d/2 \qquad \text{(E14)}$$

$$\Sigma_\infty = \frac{\omega\mu_\infty N_A D}{2} \qquad \text{(E15)}$$

it is obvious that Σ_0 is the conductance before any donor has diffused

out and Σ_∞ after all the donor has been diffused out. With these definitions (E16) becomes

$$\Sigma/\Sigma_0 = 1 + \frac{2.256\vartheta\mu_\infty\sqrt{D}}{\mu_0 d}\left(\frac{N_D{}^0}{N_A - N_D{}^0}\right)\sqrt{t} \tag{E17}$$

Calling the slope of this curve S leads to the result

$$S = \frac{2.256\vartheta\mu_\infty\sqrt{D}}{\mu_0 d}\left(\frac{N_D{}^0}{N_A - N_D{}^0}\right) \tag{E18}$$

or using (E14) and (E15)

$$D = \left(\frac{Sd\ \Sigma_0\ N_A}{2.256\vartheta\ \Sigma_\infty\ N_D{}^0}\right)^2 \tag{E19}$$

This is equivalent to equation (11.12).

GLOSSARY OF SYMBOLS

a	distance of closest approach of two ions of opposite sign
A	constant in expression for ρ in section on relaxation theory
A^-	concentration of ionized acceptors
A_0	A_η going with $\eta = 0$
A_1	A_η going with η_1
A_η	constant preceding the ηth eigenfunction in solution of the relaxation problem
A_s	coefficient of sin sx in Fourier expression for N_D
b	$q^2/2\kappa kT$, position of minimum in $g(r)$
B	constant in expression for ρ in section on relaxation theory
B^-	boron ion
$B(Si)$	un-ionized boron in silicon
B_s	coefficient of cos sx in Fourier expression for N_D
$c(r)$	concentration of positive ions in atmosphere of a negative ion
C	concentration of LiB^-
d	thickness of wafer in diffusion experiment
D	diffusivity of donor ion in the most general sense
D_0	diffusivity of donor ion in the absence of pairing
D^+	concentration of ionized donors
$D_0{}^+$	value of D^+ in the absence of acceptor
$D_*{}^+$	concentration of mobile donor ions where $V = 0$
e^-	conduction band electron
e^+	valence band hole

e^+e^-	recombined bole-electron pair
E	energy level in electron gas
E_D	ionization energy of a donor
E_A	ionization energy of an acceptor
E_i	energy level in conduction band
$E(r)$	chance that volume $4\pi r^3/3$ will not contain an ion
f	flux density
F	Fermi level — also constant in equation (7.21)
g_i	density of states of energy E_i in conduction band
$g(r)$	nearest neighbor distribution function at equilibrium
G	Gibbs free energy of electron assembly
Ga^-	gallium ion in germanium
G_η	space dependent part of relaxation eigenfunction
G_0	G_η for $\eta = 0$
G_1	G_η for $\eta = \eta_1$
h	Plank's constant — also used for normalizing constant in $c(r)$
h_j	number of holes in the jth energy level
H	net local density of fixed donors
$i(\rho_2, \rho_1)$	$\varepsilon^3 I(r_2, r_1)$
I	field current in diffusion measurement
$I(r_2, r_1)$	integral for ion pairing calculations taken between r_1 and r_2
$\vec{J}(\vec{r})$	current in the atmosphere of a nearest neighbor
J^*	flux density of ions being trapped
k	Boltzmann's constant
k_1	first order rate constant in relaxation theory
k_2	second order rate constant in relaxation theory
K_0	distribution coefficient of donor between semiconductor and external phase
K_1	electron-hole recombination equilibrium constant
K_A	ionization constant of acceptor
K_D	ionization constant of donor
K_j	constant relating ω_j to volume, V
K^*	product of K_D, K_0, and α
ℓ	screening length for diffusion potential
L	Debye length — also used for radius of volume, $1/N$
Li^+	lithium ion
$Li(Sn)$	lithium in molten tin
$Li(Si)$	un-ionized lithium in silicon
$LiSi$	lithium-silicon complex
LiB	un-ionized LiB^-

LiB^-	lithium-boron complex ion in semiconductor
$[Li^+B^-]$	lithium-boron ion pair
$[Li^+Ga^-]$	lithium-gallium ion pair
m_0	normal mass of electron
m_p	effective mass of a hole
M	normalizing constant in relaxation theory
n	concentration of conduction electrons — also used for density of untrapped ions in relaxation
n_i	intrinsic concentration of electrons
N_A	total acceptor concentration
N_D	total donor concentration
N_D^0	total solubility of donor in undoped semiconductor—also used for initial density of donors in diffusion experiments
N	ion concentration in an electrolyte solution—also used for initial value of n in relaxation—also used for concentration of immobile donors in Appendix A
N_D^*	solubility of donor in absence of ion pairing in Appendix A
p	concentration of holes
P	concentration of ion pairs
q	charge on an ion
$Q(\alpha)$	tabulated integral for computing Ω
r	distance between positive and negative ions in a pair
r_1	a particular value of r
r_2	a particular value of r
R	capture radius of an ion in relaxation
S	slope of Σ/Σ_0 versus $\sqrt{t}$ curve
S_η	time dependent part of relaxation eigenfunction belonging to eigenvalue η
t	time
T	temperature
u	$V - V_0$
V	electrostatic potential — also used for volume — also used for potential difference between probe points — also used for potential energy of a positive in neighborhood of negative ion
V_0	electrostatic potential where $x = x_0$
x	variable of integration — same as r also rectangular position coordinate
x_0	special value of x.

α — ε/a—also used for thermodynamic activity of donor in external phase

α_s — coefficient of $\sin sx$ in Fourier expression for u

β — constant in exponential in LiB^- equilibrium constant

β^* — constant in exponential in expression for vacancy concentration

β_0 — β_s for $s = 0$

β_s — coefficient of $\cos sx$ in fourier expression for u

γ — pre-exponential factor in LiB^- equilibrium constant

γ^* — pre-exponential factor in expression for vacancy concentration

γ_0 — $\exp[-qV_0/kT]$

$\Gamma(\vec{r})$ — non-equilibrium nearest neighbor distribution function

δ — constant appearing in Appendix C

ε — $q^2/\kappa kT$

r — eigenvalue in relaxation problem

η_1 — second eigenvalue in set of η

θ — fraction of donor paired

ϑ — correction factor for variable carrier mobility

κ — dielectric constant

λ — x/ε

λ_s — $2\pi/s$, wavelength of sth component of fourier series

λ_{min} — wavelength of component of fourier series for N_D, having minimum wavelength

μ — chemical potential of donor in an external phase — also used for mobility of donor ion — also used for local carrier mobility

μ^0 — chemical potential of donor in external phase in standard state

μ_{D^+} — chemical potential of donor ion

$\mu_{D^+}^0$ — chemical potential of donor ion in the standard state

μ_e — chemical potential of an electron

μ_D — chemical potential of donor atom in semiconductor

μ_D^0 — chemical potential of donor atom in standard state

μ_0 — mobility of donor atom at infinite dilution — also used for carrier mobility in diffusion experiments before diffusion

μ_∞ — carrier mobility in diffusion experiments after all diffusant bas diffused out

ν — $x/2\sqrt{Dt}$

ξ — ε/r

π — LiB^- equilibrium constant

ρ — resistivity of gallium-doped germanium after saturation with

	lithium — also used for local charge density in Poisson's equation — also used for density of diffusing positive ions in relaxation
ρ_0	resistivity of gallium-doped germanium before saturation with lithium
ρ_1	r_1/ε
ρ_2	r_2/ε
σ	conductivity during relaxation
σ_∞	conductivity in relaxed state
Σ	conductance between probe points
Σ_0	conductance before diffusion begins in diffusion experiments
Σ_∞	conductance after diffusion is over in diffusion experiments
τ	relaxation time
Φ	constant in relaxation formula for conductivity
Ψ	local electrostatic potential in ionic atmosphere
ω	proportionality constant connecting conductance between probe points with integral over carrier concentration
ω_j	number of states in jth level of electronic energy diagram
Ω	ion pairing equilibrium constant
$\square$	vacant lattice site in covalent crystal
$\square^-$	negatively charged cation vacancy

REFERENCES

1. Wagner, C., Z. Phys. Chem., **B21,** p. 25, 1933, **B32,** p. 447, 1936.
2. Taylor, H. S., and Taylor, H. A., Elementary Physical Chemistry, p. 343, Van Nostrand, 1937.
3. Shockley, W., Electrons and Holes in Semiconductors, p. 6, Van Nostrand, 1950.
4. Shockley, W., Electrons and Holes in Semiconductors, Van Nostrand, 1950.
5. Reiss, H., J. Chem. Phys., **21**, p. 1209, 1953.
6. Reiss, H., and Fuller, C. S., J. Metals, **12**, p. 276, 1956.
7. Fuller, C. S., and Ditzenberger, J. A., J. App. Phys., **25**, p. 1439, 1954.
8. Fuller, C. S., and Ditzenberger, J. A., Phys. Rev., **91**, p. 193, 1953.
9. MacDougall, F. H., Thermodynamics and Chemistry, p. 143, Wiley, 1939.
10. Miller, F. W., Elementary Theory of Qualitative Analysis, p. 102, Century Company, New York, 1929.
11. Fuller, C. S., Record of Chemical Progress, **17**, No. 2, 1956.
12. Wagner, C., and Crünewald, K., Z. Phys. Chem., **B40**, p. 455, 1938.
13. von Baumbach, H. H., and Wagner, C., Z. Phys. Chem., **22B**, p. 199, 1933.
14. Kröger, F. A., and Vink, H. J., Physica, **20**, p. 950, 1954.
15. MacDougall, F. H., Thermodynamics and Chemistry, p. 258, Wiley, 1939.
16. Shockley, W., Electrons and Holes in Semiconductors, p. 231, Van Nostrand, 1950.
17. Mayer, J. E., and Mayer, M. G., Statistical Mechanics, p. 120, Wiley, 1940.
18. MacDougall, F. H., Thermodynamics and Chemistry, p. 137, Wiley, 1939.
19. Lewis, C. N., and Randall, M. C., Thermodynamics, p. 258, McGraw Hill, 1923.

20. Mayer, J. E., and Mayer, M. G., Statistical Mechanics, p. 121, Wiley, 1940.
21. MacDougall, F. H., Thermodynamics and Chemistry, p. 261, Wiley, 1939.
22. MacDougall, F. H., Thermodynamics and Chemistry, p. 25, Wiley, 1939.
23. Engell, H. J., and Houffe, K., Z. Electrochem., **56**, p. 366, 1952, **57**, p. 762, 1953.
24. Shockley, W., Electrons and Holes in Semiconductors, p. 15, Van Nostrand, 1950.
25. Morin, F. J., and Maita, J. P., Phys. Rev., **94**, p. 1525, 1954.
26. Morin, F. J., and Maita, J. P., Phys. Rev., **96**, p. 28, 1954.
27. Shockley, W., Electrons and Holes in Semiconductors, p. 86, Van Nostrand, 1950.
28. Shockley, W., Electrons and Holes in Semiconductors, p. 88, Van Nostrand, 1950.
29. Fowler, R. H., Statistical Mechanics, p. 48, Cambridge, 1929.
30. Slater, J. C., and Frank, N. H., Introduction to Theoretical Physics, p. 212, McGraw Hill, 1933.
31. Shockley, W., B.S.T.J., **28**, p. 435, 1949.
32. Fuller, C. S., and Ditzenberger, J. A., J. App. Phys., May, 1956.
33. Shulman, R. G., and McMahon, M. E., J. App. Phys., **24**, p. 1267, 1953.
34. Reiss, H., Fuller, C. S., and Pietrusczkiewicz, A. J., J. Chem. Phys. (in press).
35. Eyring, H., Walter, J., and Kimball, G. E., Quantum Chemistry, p. 231, Wiley, 1946.
36. Pauling, L., The Nature of the Chemical Bond, p. 179, Cornell, 1942.
37. Debye, P., and Huckel, E., Phys. Z., **24**, p. 195, 1923.
38. Kirkwood, J. G., J. Chem. Phys., **2**, p. 767, 1934.
39. Briggs, H. B., Phys. Rev., **77**, p. 287, 1950.
40. Briggs, H. B., Phys. Rev., **77**, p. 287, 1950.
41. Wyman, Phys. Rev., **35**, p. 623, 1930.
42. Bjerrum, N., Kgle. Danske Vidensk. Selskab., **7**, No. 9, 1926.
43. Fuoss, R. M., Trans. Faraday Soc., **30**, p. 967, 1934.
44. Reiss, H., J. Chem. Phys. (in press).
45. Reiss, H., J. Chem. Phys. (in press).
46. Shockley, W., and Read, W. T., Jr., Phys. Rev., **87**, p. 835, 1952, Haynes, J. R., and Hornbeck, J. A., Phys. Rev., **90**, p. 152, 1953, **97**, p. 311, 1955.
47. Harned and Owen, The Physical Chemistry of Electrolytes, p. 123, A. C. S. Monograph, 1950.
48. Carslaw, H. S., and Jaeger, J. C., Conduction of Heat in Solids, p. 209, Oxford, 1948.
49. Glasstone, S., Textbook of Physical Chemistry, p. 1231, Van Nostrand, 1940.
50. Shockley, W., Electrons and Holes in Semiconductors, p. 300, Van Nostrand, 1950.
51. Slater, J. C., and Frank, N. H., Introduction to Theoretical Physics, p. 186, McGraw Hill, 1933.
52. Fuller, C. S., and Severiens, J. C., Phys. Rev., **95**, p. 21, 1954.
53. Shockley, W., B.S.T.J., **28**, p. 435, 1949.
54. Shockley, W., Electrons and Holes in Semiconductors, p. 258, Van Nostrand, 1950.
55. Theuerer, H. C., U. S. Pat. No. 2542727.
56. Margeneau, H., and Murphy, G. M., The Mathematics of Physics and Chemistry, p. 213, Van Nostrand, 1943.
57. Margeneau, H., and Murphy, G. M., The Mathematics of Physics and Chemistry, p. 72, Van Nostrand, 1943.
58. Reiss, H., and La Mer, V. K., J. Chem. Phys., **18**, p. 1, 1950.
59. Reiss, H., J. Chem. Phys., **19**, p. 482, 1951.
60. Carslaw, H. S., and Jaeger, J. C., Conduction of Heat in Solids, p. 40, Oxford, 1948.
61. Boltzmann, L., Ann. Phys., **53**, p. 959, 1894.
62. Carslaw, H. S., and Jaeger, J. C., Conduction of Heat in Solids, p. 41, Oxford, 1948.
63. Prince, M. B., Phys. Rev., **92**, p. 681, 1953, **93**, p. 1204, 1954.
64. Tyler, W. W., and Woodbury, H. H., Bull. Am. Phys. Soc., **30**, No. 7, p. 32, 1955.
65. Debye, P. P., Conwell, E. M., Phys. Rev., **93**, p. 693, 1954.

66. Shockley, W., Electrons and Holes in Semiconductors, Chapter 8, Van Nostrand, 1950.
67. Shockley, W., Electrons and Holes in Semiconductors, Chapter 16, Van Nostrand, 1950.
68. Geballe, T. H., and Morin, F. J., Phys. Rev., **95**, p. 1085, 1954.
69. Conant, J. B., The Chemistry of Organic Compounds, p. 196, Macmillan, 1939.
70. Valenta, M., and Ramasastry, C., Symposium on Semiconductors, Meeting I.M.D., and A.I.M.E., Feb. 20, 1956.
71. Longini, R. E., and Green, R., Phys. Rev. (in press).